

Development and Evaluation of a Simulation-Based Algorithm to Optimize the Planning of Interim Analyses for Clinical Trials in ALS

Jordi W.J. van Unnik, MSc, Stavros Nikolakopoulos, PhD, Marinus J.C. Eijkemans, PhD, Jésus Gonzalez-Bermejo, MD, PhD, Gaele Bruneteau, MD, PhD, Capucine Morelot-Panzini, MD, PhD, Leonard H. van den Berg, MD, PhD, Merit E. Cudkowicz, MD, Christopher J. McDermott, MD, PhD, Thomas Similowski, MD, PhD, and Ruben P.A. van Eijk, MD, PhD, for the DiPALS, RespiStimALS and Ceftriaxone Study Group

Correspondence

Dr. van Eijk
r.p.a.vaneijk-2@umcutrecht.nl

Neurology® 2023;100:e2398-e2408. doi:10.1212/WNL.0000000000207306

Abstract

Background and Objectives

Late-phase clinical trials for neurodegenerative diseases have a low probability of success. In this study, we introduce an algorithm that optimizes the planning of interim analyses for clinical trials in amyotrophic lateral sclerosis (ALS) to better use the time and resources available and minimize the exposure of patients to ineffective or harmful drugs.

Methods

A simulation-based algorithm was developed to determine the optimal interim analysis scheme by integrating prior knowledge about the success rate of ALS clinical trials with drug-specific information obtained in early-phase studies. Interim analysis schemes were optimized by varying the number and timing of interim analyses, together with their decision rules about when to stop a trial. The algorithm was applied retrospectively to 3 clinical trials that investigated the efficacy of diaphragm pacing or ceftriaxone on survival in patients with ALS. Outcomes were additionally compared with conventional interim designs.

Results

We evaluated 183–1,351 unique interim analysis schemes for each trial. Application of the optimal designs correctly established lack of efficacy, would have concluded all studies 1.2–19.4 months earlier (reduction of 4.6%–57.7% in trial duration), and could have reduced the number of randomized patients by 1.7%–58.1%. By means of simulation, we illustrate the efficiency for other treatment scenarios. The optimized interim analysis schemes outperformed conventional interim designs in most scenarios.

Discussion

Our algorithm uses prior knowledge to determine the uncertainty of the expected treatment effect in ALS clinical trials and optimizes the planning of interim analyses. Improving futility monitoring in ALS could minimize the exposure of patients to ineffective or harmful treatments and result in significant ethical and efficiency gains.

From the Department of Neurology (J.W.J.v.U., L.H.v.d.B., R.P.A.v.E.), UMC Utrecht Brain Center, and Biostatistics & Research Support (S.N., M.J.C.E., R.P.A.v.E.), Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, the Netherlands; Sorbonne Université (J.G.-B., C.M.-P., T.S.), INSERM, UMR51158 Neurophysiologie Respiratoire Expérimentale et Clinique, AP-HP, Groupe Hospitalier Universitaire APHP-Sorbonne Université, Site Pitié-Salpêtrière, Département R35; APHP (G.B.), Groupe Hospitalier Paris 6, Hôpital Pitié-Salpêtrière, Département de Neurologie, Centre Référent SLA, France; Department of Neurology (M.E.C.), Massachusetts General Hospital, Boston; and Sheffield Institute for Translational Neuroscience (C.J.M.), University of Sheffield, United Kingdom.

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.

The Article Processing Charge was funded by Wolters Kluwer/UKB VSNU Agreement.

Coinvestigators are listed at links.lww.com/WNL/C773.

This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Glossary

ALS = amyotrophic lateral sclerosis; **ALSFRS-R** = Revised ALS Function Rating Scale; **DMEC** = Data Monitoring and Ethics Committee; **HR** = hazard ratio; **NIV** = noninvasive ventilation.

Drug development for amyotrophic lateral sclerosis (ALS) has been proven to be difficult. More than 70 therapeutic compounds have been evaluated over the past 30 years,¹⁻³ resulting in only minimal gains in life expectancy⁴ or amelioration of progression rates.^{5,6} Significantly, several interventions have led to an accelerated loss of function and reduction of survival time,⁷⁻¹¹ emphasizing that trial participation may not always be free from harm. The high futility rate among ALS clinical trials, together with the risks associated for patients, stresses the importance—both ethically and economically—of developing more efficient strategies, to identify ineffective or harmful treatments more quickly and thereby improve the safety of trial participants and minimize the loss of time, funding, and resources.

Currently, most ALS clinical trials perform their statistical analyses after all patients have completed a fixed follow-up period.^{12,13} As an alternative, interim analyses could be performed on the accumulating data when only a subset of patients have completed their follow-up.¹⁴ This presents the opportunity to stop a trial before all patients have been enrolled, or completed the study, as soon as there is sufficient evidence for a drug's (in)efficacy. The probability that a trial can be stopped early, however, depends strongly on the timing and the prespecified decision rules or stopping criteria.¹⁵ These criteria are often defined arbitrarily at the design stage of a study and frequently ignore information that is available from earlier clinical trials.¹⁶ As a result, the planning and stopping criteria of interim analyses may be suboptimal, which could potentially mean that ineffective treatments are continued or exposure to placebo is prolonged unnecessarily.¹²

In this study, we develop a simulation-based algorithm that determines the optimal timing of interim analyses, together with their decision rules, by incorporating information obtained in earlier stages of development, and minimize the expected trial duration or sample size. By applying the algorithm to 3 completed clinical trials, we aim to evaluate its performance and obtain insights into optimization strategies to better tailor study designs for future clinical trials.

Methods

This study consisted of 2 parts: first, we developed a simulation-based algorithm to determine the optimal timing and decision rules for interim analyses, on the basis of prior knowledge. Second, we applied the algorithm retrospectively to 3 clinical trials that investigated the efficacy and safety of diaphragm pacing or ceftriaxone in patients with ALS. We evaluated whether the studies could have been stopped

earlier, and when, and compared the results with other commonly used interim analysis schemes or designs without interim analyses.

Interim Analyses for Clinical Trials

During an interim analysis, using the accumulating data, one determines whether there is sufficient evidence to stop the study early for either drug efficacy (superiority) or lack of it (futility) or whether it is better to continue the trial and collect more information. If interim analyses are used, the investigator must prespecify: (1) the timing of the analyses (i.e., after how many months/patients/events) and (2) the decision rules about when a treatment is considered (in)effective. These decision rules are statistical criteria that define how large the (standardized) effect size or how small the p value should be to stop the trial. As such, interim analyses result in multiple statistical tests, requiring a correction of the p value for the treatment effect. There are several ways to adjust p values, for example, by using a Pocock design¹⁷ that sets the same significance threshold for all interim analyses uniformly (e.g., stop if $p < 0.01$). Or, alternatively, one can use the more common O'Brien-Fleming design¹⁸ that uses stringent significance thresholds at early interim analyses (e.g., stop if $p < 0.001$), but sets more lenient significance thresholds at later interim analyses (e.g., stop if $p < 0.045$). The benefit of the Pocock design is that it becomes easier to stop at early interim analyses and, on average, results in shorter studies. The downside is that it becomes more difficult to stop at later interim analyses if the trial cannot be stopped early. Consequently, such an interim analysis scheme potentially requires more patients or a longer duration compared with an O'Brien-Fleming design. As such, prespecifying the interim analyses at the design stage of any trial is crucial. By varying the number of interim analyses, together with their timing and decision rules, many unique interim analysis schemes can be defined with different probabilities of stopping early and expected trial durations. The scheme that one will consider to be most efficient depends primarily on the treatment effect that one expects.¹⁶

Uncertainty About the Expected Treatment Effect

The expected treatment effect, therefore, is a decisive assumption at the design stage of any trial, which will drive the efficiency of an interim analysis scheme. The expected treatment effect is usually based on either drug-specific information obtained in earlier development stages (e.g., small phase 2 studies) or defined arbitrarily. In many instances, however, there is a considerable degree of uncertainty around the expected treatment effect.¹⁹ In addition, especially for ALS, the success rate of any clinical development program is

low, and overall, most trials result in a futile conclusion.²⁰ Ideally, the probability of success and the overall uncertainty about the treatment effect need to be accounted for to find the optimal interim analysis scheme, given the information available at the design stage.

Optimization Algorithm for Interim Analyses

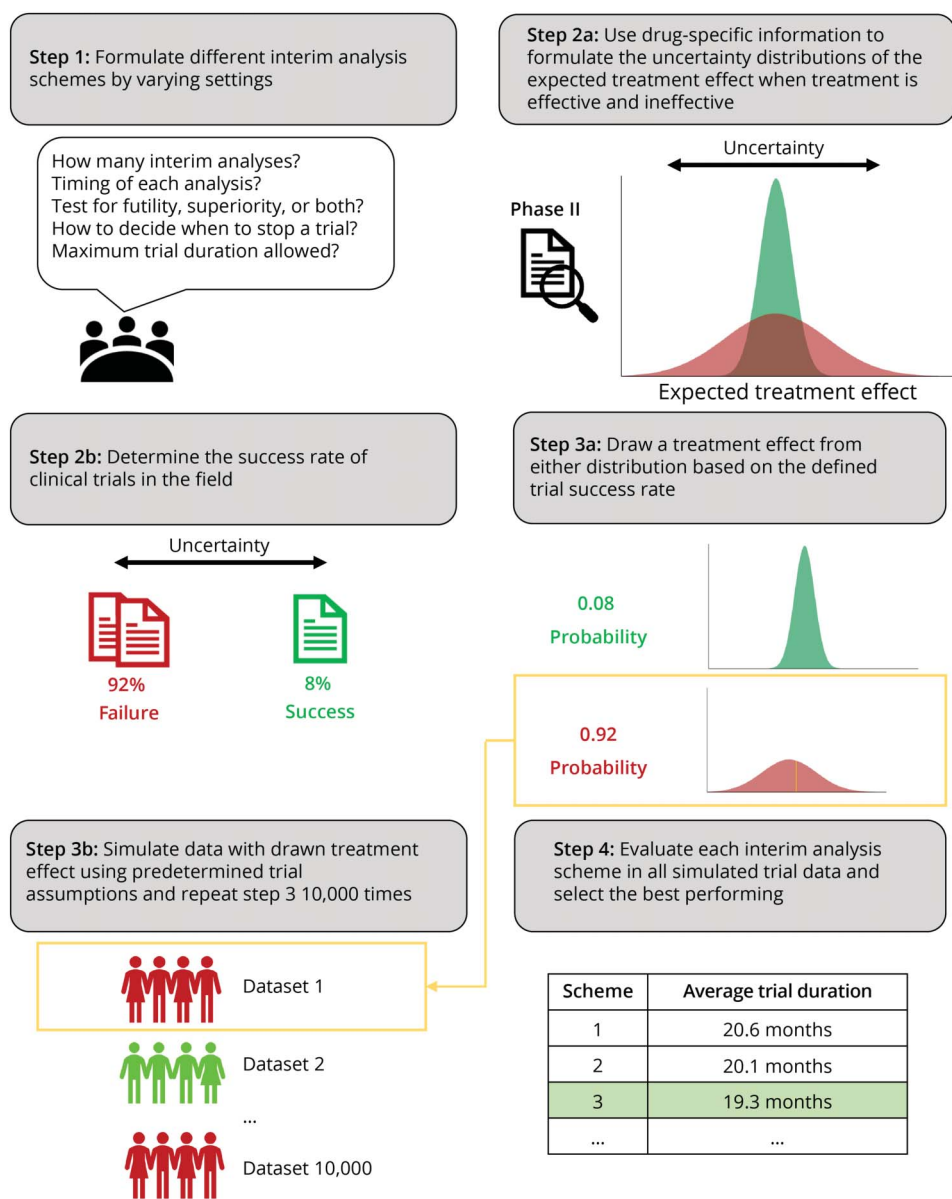
We approached this challenge by means of simulation, thereby resampling the treatment effect from a distribution reflective of both the uncertainty in the treatment effect and the overall probability of success. For each simulation, we compared different interim analysis schemes and selected the design with the most efficient operating characteristics (e.g., trial duration, randomized patients, or drug exposure). Figure 1 provides a schematic of our optimization process. First, we defined all possible

interim analysis schemes by varying the timing, number of analyses, and decision rules (step 1). Second, we defined 2 uncertainty distributions for the treatment effect: (1) for the scenario that treatment is effective and (2) for the scenario that treatment is ineffective (step 2a). Next, we defined the probability that a trial will be successful (step 2b). Finally, we sampled a treatment effect from our uncertainty distribution, conditional on whether the trial would be successful or not (step 3a). The drawn treatment effect was subsequently used to simulate a clinical trial dataset (step 3b), in which the different interim analysis schemes were evaluated (Figure 1, step 4).

Application of Algorithm to ALS Clinical Trials

We applied the abovementioned algorithm retrospectively to 3 completed ALS clinical trials (DiPALS, RespiStimALS, and

Figure 1 Schematic Illustration of the Simulation-Based Algorithm to Optimize Interim Analysis Schemes



First, the investigators decide on the range of acceptable interim analysis schemes. Previous research is then used to formulate the uncertainty around the treatment effect and the probability of trial success. In the next step, each interim analysis scheme is evaluated by means of simulation based on the trial assumptions made. Finally, the best performing interim analysis scheme is identified, based on a selection criterion (e.g., shortest average trial duration, smallest average sample size, or shortest average drug exposure). A script to replicate the simulation is provided in eAppendix 1 (links.lww.com/WNL/C768).

Table 1 Overview of the Original Trial Design Settings

	DiPALS (2015)	RespiStimALS (2016)	Ceftriaxone-ALS (2014)
Primary endpoint	Time to death	Time to NIV or death	Time to PAV, tracheostomy, or death
Expected enrollment period, mo	18	36	24
Minimum follow-up duration, mo	12	24	12
Required sample size	108	74	500
Expected survival control arm at endpoint, %	45	2.5	75
Expected survival treatment arm at endpoint, %	70	15	82.5
Expected treatment effect (HR)	0.45	0.50	0.67
Uncertainty around treatment effect	FDA report ²²	FDA report ²²	Phase II trial ²⁵
No. of planned interim analyses	0	0	5

Abbreviations: ALS = amyotrophic lateral sclerosis; FDA = Food and Drug Administration; HR = hazard ratio; NIV = noninvasive ventilation; PAV = permanent-assisted ventilation.

Ceftriaxone-ALS) to evaluate whether using an optimized interim analysis scheme could have concluded the studies earlier. The study population and methods of the trials are described elsewhere.^{10,11,21} In short, all trials were randomized controlled clinical trials that assessed the efficacy and safety of diaphragm pacing (DiPALS and RespiStimALS) or ceftriaxone (Ceftriaxone-ALS) in patients with ALS. Patients in the DiPALS study were randomized to receive either noninvasive ventilation (NIV) plus diaphragm pacing or NIV alone, whereas patients in the RespiStimALS study received either diaphragm or sham pacing. All patients receiving (sham) pacing were operated laparoscopically. The primary endpoint of the DiPALS study was time to death, whereas the RespiStimALS study used time to death or NIV. For the Ceftriaxone-ALS study, patients were randomized to receive either intravenous ceftriaxone or placebo. This study used a coprimary endpoint assessing both time to death or respiratory insufficiency and the change in daily function (Revised ALS Function Rating Scale [ALSFRS-R]). To harmonize the 3 clinical trials, we used only the survival endpoint of the Ceftriaxone-ALS study. While DiPALS and RespiStimALS studies were originally designed without interim analyses, the Ceftriaxone-ALS trial had already planned 5 interim analyses. The DiPALS and RespiStimALS trials were stopped prematurely because of safety issues resulting in excess deaths in the pacing arms, whereas the Ceftriaxone-ALS trial was stopped early for ineffectiveness. All patients provided informed consent, and all studies were approved by an appropriate ethical committee.

Both pacing trials were designed based on an unpublished, historically controlled, multicenter cohort study that led to Humanitarian Device Exemption approval of diaphragm pacing by the US Food and Drug Administration.²² In that report, the median survival of the historical cohort who received usual care was 21.4 months ($n = 43$), whereas the median survival in the cohort who received pacing was 37.5

months ($n = 43$). To define the uncertainty around the expected treatment effect, we assumed that the cohort was followed up for 24 months and that survival time followed a Weibull distribution ($\rho = 2$),²³ resulting in a standard error of the log-hazard ratio (HR) of 0.335.²⁴ Significantly, the success rate of any phase 3 clinical trial in ALS is 7.1% (2 successes of 28 trials).²⁰ As such, one should be even more uncertain about the expected treatment effect. To inflate the uncertainty around the expected treatment effect, we defined a mixture distribution, sampling the standard error of the log-HR from a β distribution with a 0.071 probability of 0.335 and a $1 - 0.071$ probability of 1.06 (i.e., 10 times the variance of the historical data). To illustrate, if we were to design a trial with a HR of 0.45, the 75% uncertainty level around that estimate, conditional on the success probability and historical data, ranges from 0.23 to 0.87. For the ceftriaxone trial, we assumed a median survival of 28.9 months for patients receiving placebo and a median survival of 43.4 months for those receiving ceftriaxone. Based on the observed functional decline of the ALSFRS-R in the phase II trial,²⁵ and assuming that every point increase in ALSFRS-R reduces the hazard for death by 11.9%,²⁶ we estimated the standard error of the log-HR for ceftriaxone to be 0.147. We then defined the mixture distribution similar as described earlier. Based on the assumptions defined in the respective trial protocol (Table 1), we simulated clinical trial data for each study and determined their optimal interim analysis scheme.

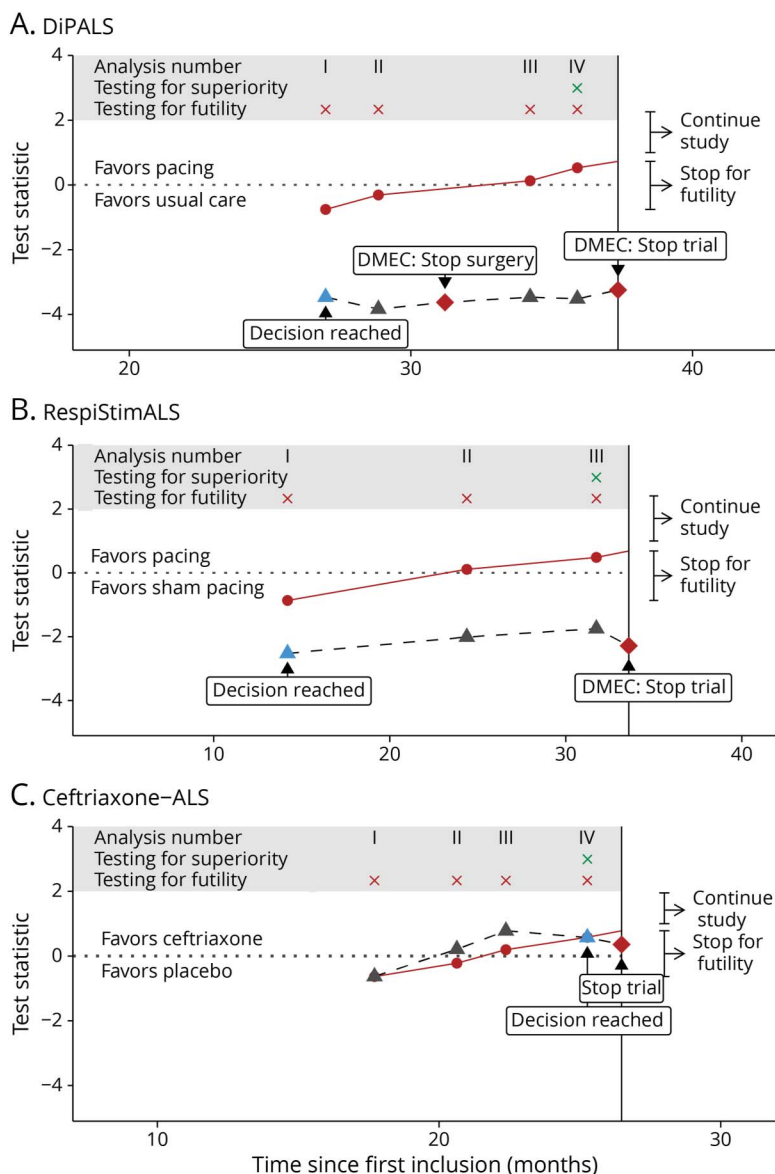
Statistical Analysis

Time-to-event data were simulated using a Weibull distribution assuming a constant enrollment rate and no loss to follow-up for the primary endpoint. A log-rank test was used to compare treatment arms after a certain number of events had been reached (i.e., an event-driven trial design).²⁷ In each simulated dataset, we evaluated a variety of interim analysis schemes. These varied in number of futility analyses from 4 to 5 and in timing ranging from 30% to 90% of the maximum

number of events, with steps of 10%. In all schemes, the interim analysis for superiority was fixed at 60% because stopping early for efficacy is often undesirable.^{28,29} The nonbinding futility boundary ranged from Pocock type to O'Brien-Fleming type (i.e., β -spending function from 0.75 to 3.00 with steps of 0.25), whereas the superiority boundary ranged from 2.00 to 3.00.³⁰ Schemes that had a maximum trial duration of more than 10% compared with a design without interim analyses were excluded.³¹ Clinical trial scenarios were simulated 10,000 times. The interim testing scheme with the shortest average trial duration was selected as the most optimal design. In addition, we applied 4 conventional interim designs to the 3 completed ALS clinical trials to compare the performance of the optimized interim analysis schemes to other design approaches. The number of futility analyses in these conventional interim designs varied from 1 to 4 at

arbitrarily defined time points using O'Brien-Fleming-type decision rules.¹⁸ Finally, to investigate alternative treatment effects, we simulated the treatment effect, from very beneficial to very harmful, in each of 3 trials while keeping enrollment rate, survival in the placebo arm, and randomization ratio fixed. For each scenario, we determined when the trial could be stopped, using designs either with or without (optimized) interim analyses. We highlight the comparison of the optimized interim analysis schemes with designs without interim analyses or with a single interim analysis at 60% of the maximum number of events, which reflects commonly used interim design settings.³² All statistical analyses were performed with the R language for statistical programming and the package *Rpact* (version 3.2.2; Wassmer G and Pahlke F, 2022); the script to replicate the simulations is provided in eAppendix 1 (links.lww.com/WNL/C768).

Figure 2 Retrospective Application of the Algorithm in the 3 ALS Clinical Trials



Performance of the optimized interim analysis schemes when applied retrospectively to the DiPALS (A), RespiStimALS (B), and Ceftriaxone-ALS (C) trials. The black dashed line reflects the development of the test statistic over time, which is calculated sequentially and represents the treatment evidence based on the data accumulated thus far. As soon as the treatment evidence surpasses the red line, the trial can be stopped for futility. This line, therefore, represents the decision rules of the trial. In addition, a horizontal dotted line was plotted to indicate whether treatment evidence favors treatment or control. The blue triangle represents the stopping decision of the optimized interim analysis scheme, whereas the red diamond reflects when the trial was actually stopped. ALS = amyotrophic lateral sclerosis; DMEC = Data Monitoring and Ethics Committee.

Data Availability

Investigators may request access to the de-identified individual participant data from the respective corresponding authors of the DiPALS and RespiStimALS studies or from the National Institute of Neurological Disorders and Stroke for access to the Ceftriaxone-ALS data.

Results

Based on information available at the design stage, we evaluated 797 interim analysis schemes for the DiPALS trial, 183 schemes for the RespiStimALS trial, and 1,351 schemes for the Ceftriaxone-ALS trial. The number of eligible schemes varied between trials due to differences in the original design settings and their implications for the maximum trial duration. For the DiPALS trial, the optimal interim analysis scheme consisted of 4 interim analyses at 30%, 40%, 50%, and 60% of the (maximum) planned number of events (α -spending and β -spending parameters: 2.50). For the RespiStimALS trial, the optimal scheme consisted of 4 interim analyses at 30%, 50%, 60%, and 70% (α -spending and β -spending parameters of 2.25 and 2.75, respectively). For the Ceftriaxone-ALS trial, the optimal scheme consisted of 4 interim analyses at 30%, 40%, 50%, and 60% (α -spending and β -spending parameters: 2.25).

Application of Optimized Interim Analysis Scheme

The DiPALS trial was discontinued prematurely by the Data Monitoring and Ethics Committee (DMEC). The stop decision is illustrated in Figure 2A together with the test statistic for the treatment effect at each time point. In addition, we provide the decision boundaries for the optimized interim analysis scheme. As can be seen, the observed test statistic for the treatment effect at the first interim analysis is lower than the futility boundary. As such, the trial could have been stopped for futility at that time, leading to reduction of 34.0% in trial duration, and could have reduced the number of randomized patients by 17.6%.

A comparison with the original design is detailed in Table 2 and with a nonoptimized interim analysis scheme in Table 3. In Figure 3, we illustrate the maturity of the survival curve at

each interim analysis for the DiPALS study, together with the probability of reaching statistical significance if the trial had continued, conditional on the expected treatment effect at the design stage. For example, at the first interim analysis, we observed an HR of 2.15 in favor of usual care. The probability, if we were to continue the trial and collect more data, of this treatment effect shifting from an HR of 2.15 toward statistical significance in favor of treatment, is 6.3%. As can be seen, this so-called “conditional power” approaches zero at subsequent interim analyses, effectively ruling out the possibility that the trial will be successful.

Similarly, Figure 2, B and C, presents the decisions for the RespiStimALS and Ceftriaxone-ALS trials. In both cases, the optimized interim analysis scheme could reach a conclusion before the original design, leading to a reduction in trial duration of 57.7% for RespiStimALS and of 4.6% for Ceftriaxone-ALS, and could have reduced the number of randomized patients by 58.1% and 1.7%, respectively (Table 2). In both cases, a conventional interim design would be less efficient (Table 3). For other conventional interim designs, we provide an overview of the results for each trial in eTable 1 (links.lww.com/WNL/C770). In all 3 trials, each interim analysis scheme would outperform a design without interim analyses.

Finally, in Figure 4, we present the behavior of the optimized interim analysis scheme under different treatment effects for ceftriaxone. Likewise, we report the results for diaphragm pacing in eFigure 1 (links.lww.com/WNL/C769). As can be seen, the optimized scheme is most effective when there is a large positive or negative treatment effect in reducing trial duration when compared with a design without interim analyses. In eTable 2 (links.lww.com/WNL/C771), we report the results of other conventional interim designs had the trial concluded a neutral treatment effect (i.e., HR = 1). For each trial, the optimized interim scheme would, on average, be the most efficient design.

Discussion

In this study, we have explored an optimization strategy for the planning of interim analyses in clinical trials. As opposed to using arbitrary design settings, our simulation-based

Table 2 Outcomes of Retrospective Application of Optimized Interim Analysis Schemes in 3 ALS Clinical Trials

Outcomes	DiPALS		RespiStimALS		Ceftriaxone-ALS	
	Original design	Optimized ^a	Original design	Optimized ^a	Original design ^b	Optimized ^a
Patients randomized	74	61 (–17.6%)	74	31 (–58.1%)	461	453 (–1.7%)
Patients treated	32	25 (–21.9%)	73	31 (–57.5%)	303	298 (–1.7%)
Cumulative treatment exposure, mo	323	181 (–44.0%)	215	71 (–67.1%)	4,208	3,980 (–5.4%)
Stop trial	June 23, 2014	August 10, 2013	July 16, 2015	December 3, 2013	August 19, 2011	July 13, 2011
Trial duration, mo	30.6	20.2 (–34.0%)	33.6	14.2 (–57.7%)	26.5	25.3 (–4.6%)

Abbreviation: ALS = amyotrophic lateral sclerosis.

^a Optimized interim analysis scheme following the simulation-based algorithm described in Figure 1.

^b Original interim analysis scheme following trial design settings described in Table 1.

Table 3 Outcomes of Retrospective Application of Conventional Interim Designs in 3 ALS Clinical Trials

Outcomes	DiPALS		RespiStimALS		Ceftriaxone-ALS	
	Conventional ^a	Optimized ^b	Conventional ^a	Optimized ^b	Conventional ^a	Optimized ^b
Patients randomized	74 (−0.0%)	61 (−17.6%)	71 (−4.1%)	31 (−58.1%)	513 (+11.3%)	453 (−1.7%)
Patients treated	32 (−0.0%)	25 (−21.9%)	70 (−4.1%)	31 (−57.5%)	340 (+12.2%)	298 (−1.7%)
Cumulative treatment exposure, mo	308 (−4.6%)	181 (−44.0%)	193 (−10.1%)	71 (−67.1%)	5,942 (+41.2%)	3,980 (−5.4%)
Stop trial	May 9, 2014	August 10, 2013	May 11, 2015	December 3, 2013	May 29, 2012	July 13, 2011
Trial duration, mo	29.1 (−4.8%)	20.2 (−34.0%)	31.4 (−6.5%)	14.2 (−57.7%)	35.8 (+35.2%)	25.3 (−4.6%)

Abbreviation: ALS = amyotrophic lateral sclerosis.

^a Conventional interim design incorporated a single interim analysis at 60% of the maximum number of events using O'Brien-Fleming-type decision rules.¹⁸

^b Optimized interim analysis scheme following the simulation-based algorithm described in Figure 1.

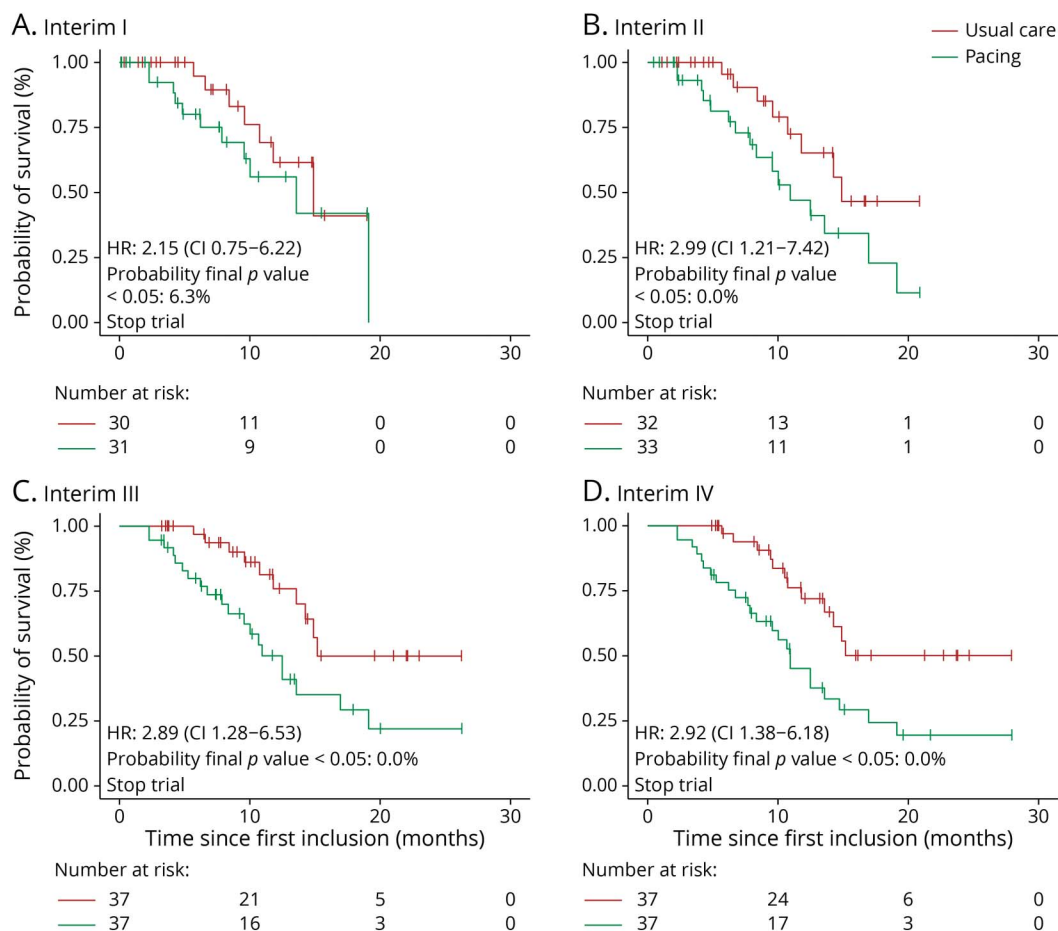
algorithm uses information from earlier clinical trials in combination with historical success probabilities to optimize the study design. Objective selection of the most optimal monitoring scheme can then be based on the preferred operational characteristics, such as the expected trial duration, number of patients, or cumulative drug exposure time. By retrospective application of our algorithm to 3 completed ALS trials, we have shown that optimized interim analysis schemes may lead to considerable reductions in trial duration and number of patients randomized when compared with designs without interim analyses or with arbitrarily defined interim analysis schemes. Consequently, the time of patient exposure to ineffective or harmful treatments can be minimized, and resources may be reallocated sooner to other promising therapeutic candidates.

Considering the high futility rates in ALS clinical trials, the pressing unmet medical need, and the ever-increasing pipeline of potential pharmaceutical agents,³ it is imperative to identify (in)effective or harmful treatments quickly. This is further underlined by the conduct of several trials in the past, which have led to accelerated loss of function and reduction in survival time compared with usual care.^{7–11} Interim analyses have been widely reported to be beneficial,³³ a conclusion supported by our findings. Although the use of such analyses has been the subject of requests from the patient community and consensus guidelines,^{34,35} surprisingly, they have not been implemented very frequently in ALS trials.^{12,13} Common barriers may include lack of expertise and experience, increased effort required in the planning and conduct of analyses, insufficient funding to compensate for lack of design familiarity, and the preference for more familiar methods.³⁶ Even when interim analyses are implemented, their planning and decision rules are often based on arbitrary and subjective choices. This is undesirable because the timing of interim analyses, and their stopping rules, have a significant impact on the study efficiency and probability to stop early.¹⁶ Our algorithm resolves this challenge by enabling an objective and data-driven approach to the planning of interim analyses and their decision rules, while not increasing the complexity of the study.

The optimal interim analysis scheme depends primarily on the treatment effect that one expects to observe.¹⁶ The uncertainty around the expected treatment effect is typically assumed to be zero, which may affect the trial if the actual treatment effect is considerably different. This risk can be mitigated by formulating an uncertainty distribution around the expected treatment effect.¹⁹ For example, if there is no evidence of any therapeutic benefit at the design stage, or if the success rate of clinical development is low, one may want to focus primarily on futility, while one would be more focused on stopping early for efficacy if a previous trial had been successful. Our algorithm automates such considerations and assesses the ideal trade-off between early stopping for futility and/or efficacy based on the uncertainty in the treatment effect at the design stage. In some settings, however, stopping early for efficacy could be undesirable, and limiting the number of superiority analyses may be warranted.^{28,29} We have accommodated this by implementing asymmetric testing schemes that solely test for futility at early interim analyses and only evaluate efficacy after a certain amount of information has been collected.

The amount of information required should be evaluated on a case-by-case basis and depends on the development phase of the trial. One consideration, for example, includes whether the results would be sufficiently persuasive for regulators if the trial were stopped at the interim analysis. Moreover, in some settings, aggressive futility monitoring carries the risk of terminating a trial early that might have demonstrated a beneficial effect had it been continued. This risk, however, does not differ from a design without futility monitoring because these error probabilities are fixed from the outset and equal to a design without interim analyses. An important exception to the rule includes a study in which treatment has a delayed effect on the outcome.³⁷ In such instances, very early futility monitoring may increase the possibility of drawing the wrong conclusion. When planning early interim analyses, therefore, the trade-off between early stopping risks, the gain in efficiency, and the minimum required level of information should be considered carefully.

Figure 3 Development of the Primary Outcome in the DiPALS Trial



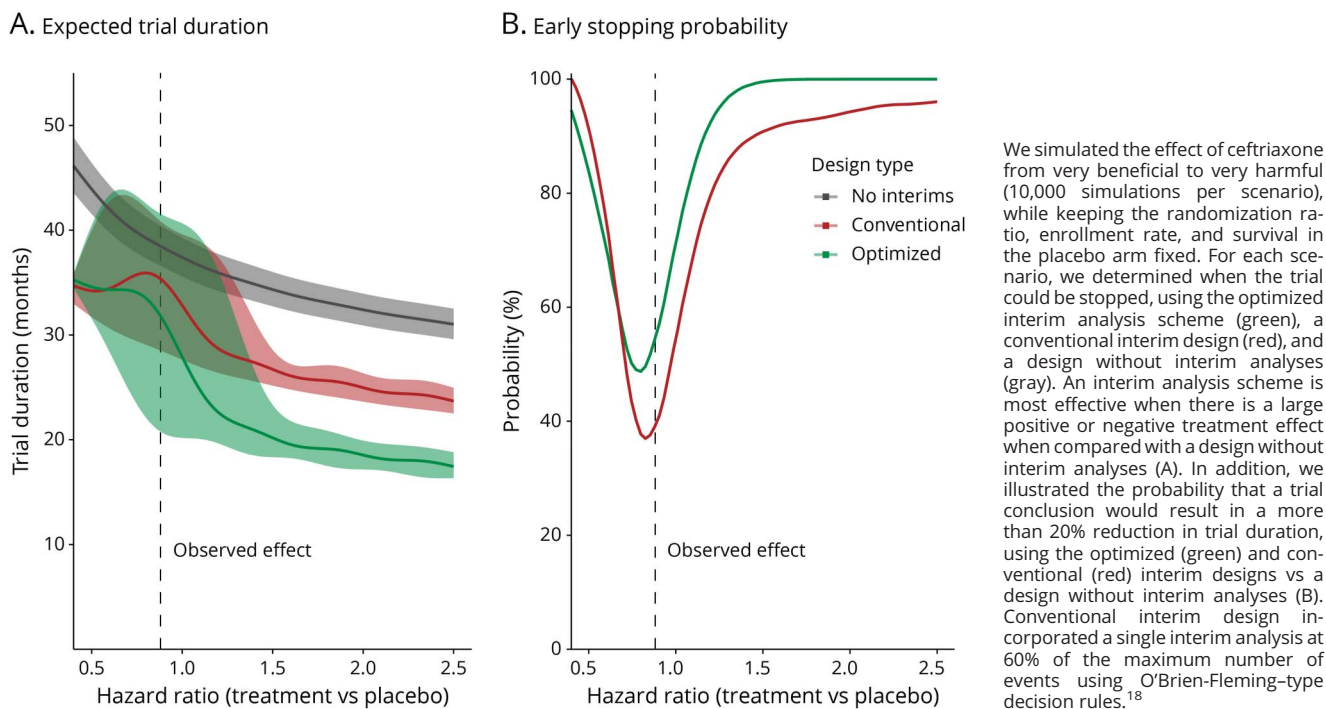
Development of the observed treatment effect over time at each interim analysis (A–D). For illustrative purposes, we show the development of the Kaplan-Meier curves at each interim analysis of the DiPALS trial. At the first interim analysis (A), the hazard ratio (HR) is 2.15 (0.75–6.22), and the trial can be stopped (Figure 2A). Taking into account the accumulated data at this time, and the expected HR of 0.45, the probability that the final analysis may still yield statistical significance in favor of treatment is only 6.3%.

In our study, we considered only early trial termination. However, additional strategies may be of interest to further advance the benefit of interim analyses, such as sample size reestimation, adjustment of the randomization ratio, or enriching the study through amended eligibility criteria. Others have shown how such modifications can positively affect the conduct and efficiency of clinical trials clinical trials^{38,39} and could be valuable extensions of our proposed algorithm. Especially, the modification of eligibility criteria could be of particular relevance for a heterogeneous disease such as ALS, where some patients could benefit more from treatment than others.⁴⁰ Prospectively implementing interim analyses to enrich the trial population may help to identify responding subgroups earlier and may overcome the need to run additional confirmative studies. These efforts may be further improved by using historical data during the interim analysis, for example, within a Bayesian framework, and borrow information to reduce sample sizes or improve precision,^{41,42} or making better use of prediction rules to improve the identification of subgroups.^{43,44} Of importance, these endeavors should be further expanded to key

intermediate outcomes, such as the ALSFRS-R and vital capacity, to increase generalizability of our findings, for example, by retrospective application in clinical trials with beneficial effects, such as edaravone, sodium phenylbutyrate-taurursodiol, or methylcobalim.^{5,45,46}

Our study has a few limitations. First, we assumed that an interim analysis would be conducted as soon as the target number of events is reached. This is, however, overoptimistic because there is usually a delay between the date the event is reached and the actual analysis. In practice, therefore, the efficiency of the interim analysis scheme could be maximized by ensuring efficient communication between trial staff and the DMEC and by optimizing data flow.⁴⁷ Nevertheless, the amount of time required to complete an analysis may remain considerable, meaning it is not feasible to conduct 4 or more interim analyses for some settings. The algorithm, however, can be easily adjusted to limit the candidate interim analysis schemes to solely 2 or 3 analyses. Our algorithm may be further improved by also considering other sources of uncertainty, such as the uncertainty in the hazard rate, expected

Figure 4 Behavior of the Optimized Interim Analysis Scheme Under Different Treatment Effects for Ceftriaxone



enrollment rate, and dropout percentage. Although these variables have no impact on the statistical power if an event-driven design is used,²⁷ they do affect the expected trial duration and could result in different optimal monitoring schemes. Open-access initiatives, such as Answer ALS and Pooled Resource Open-Access ALS Clinical Trials, could act as key source to model these sources of uncertainty in the natural history of ALS.^{48,49}

In conclusion, the pressing unmet medical need and high failure rate of clinical trials in ALS stresses the urgency—both ethically and economically—to reform the design of studies and explore more efficient alternatives. We have proposed a strategy that aims to tailor the study design based on the available information, a priori, to not only discard futile and harmful treatments early, but also stop the study soon if there is sufficient evidence for effectiveness. The application of our optimization strategy in future ALS clinical trials may, therefore, not only minimize the exposure of patients to ineffective or harmful drugs, but also accelerate the search for beneficial treatments for this devastating disease.

Acknowledgment

The authors thank Brenda Vollers-King for assistance with language editing and the National Institute of Neurological Disorders and Stroke and David A. Schoenfeld for providing access to the de-identified individual participant data of the Ceftriaxone-ALS trial.

Study Funding

This study was funded by Stichting ALS Nederland (TRICALS-Reactive II). The DiPALS study was funded by the National

Institute for Health Research Health Technology Assessment program (09/55/33) and by the Motor Neurone Disease Association of England, Wales, and Northern Ireland. The RespiStimALS study was funded by the Hospital Program for Clinical Research, French Ministry of Health (grant no. P110133), and by a “Contrat de Recherche Clinique” of the Direction de la Recherche Clinique et du Développement (DRCD), Assistance Publique-Hôpitaux de Paris, Paris, France (grant no. CRC15017-R02), which also served as the study sponsor. It was also funded by the French patients’ association for ALS research (“Association pour la Recherche sur la Sclérose Latérale Amyotrophique,” ARSLA) and by the Thierry de Latran foundation for Amyotrophic Lateral Sclerosis. The Ceftriaxone-ALS study was funded by the National Institute of Neurological Disorders and Stroke (grant 5 U01-NS-049640).

Disclosure

J.W.J. van Unnik, S. Nikolakopoulos, M.J.C. Eijkemans, J. Gonzalez-Bermejo, G. Bruneteau, C. Morelot-Panzini, L.H. van den Berg, and M.E. Cudkowicz report no disclosures relevant to the manuscript. C.J. McDermott is supported by the NIHR Sheffield Biomedical Research Center and the NIHR Research Professorship Award. T. Similowski and R.P.A. van Eijk report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* November 18, 2022. Accepted in final form March 3, 2023. Submitted and externally peer reviewed. The handling editor was Associate Editor Anthony Amato, MD, FAAN.

Appendix 1 Authors

Name	Location	Contribution
Jordi W.J. van Unnik, MSc	UMC Utrecht, Utrecht, the Netherlands	Designed and conceptualized study; analyzed or interpreted the data; and drafted the article for intellectual content
Stavros Nikolakopoulos, PhD	UMC Utrecht, Utrecht, the Netherlands	Designed and conceptualized study; analyzed or interpreted the data; and revised the article for intellectual content
Marinus J.C. Eijkemans, PhD	UMC Utrecht, Utrecht, the Netherlands	Designed and conceptualized study; analyzed or interpreted the data; and revised the article for intellectual content
Jésus Gonzalez-Bermejo, MD, PhD	Sorbonne Université, Paris, France	Major role in the acquisition of data; revised the article for intellectual content
Gaelle Bruneteau, MD, PhD	APHP, Paris, France	Major role in the acquisition of data; revised the article for intellectual content
Capucine Morelot-Panzini, MD, PhD	Sorbonne Université, Paris, France	Major role in the acquisition of data; revised the article for intellectual content
Leonard H. van den Berg, MD, PhD	UMC Utrecht, Utrecht, the Netherlands	Designed and conceptualized study; analyzed or interpreted the data; and revised the article for intellectual content
Merit E. Cudkowicz, MD	Massachusetts General Hospital, Boston	Major role in the acquisition of data; revised the article for intellectual content
Christopher J. McDermott, MD, PhD	Sheffield Institute for Translational Neuroscience, United Kingdom	Major role in the acquisition of data; revised the article for intellectual content
Thomas Similowski, MD, PhD	Sorbonne Université, Paris, France	Major role in the acquisition of data; revised the article for intellectual content
Ruben P.A. van Eijk, MD, PhD	UMC Utrecht, Utrecht, the Netherlands	Designed and conceptualized study; analyzed or interpreted the data; revised the article for intellectual content

Appendix 2 Coinvestigators

Coinvestigators are listed at links.lww.com/WNL/C773.

References

- 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(11):1127-1138. doi:10.1016/S1474-4422(14)70129-2
- Petrov D, Mansfield C, Mousy A, Hermine O. ALS clinical trials review: 20 years of failure. Are we any closer to registering a new treatment? *Front Aging Neurosci*. 2017;9:68. doi:10.3389/fnagi.2017.00068
- Wobst HJ, Mack KL, Brown DG, Brandon NJ, Shorter J. The clinical trial landscape in amyotrophic lateral sclerosis: past, present, and future. *Med Res Rev*. 2020;40(4):1352-1384. doi:10.1002/med.21661
- Lacomblez L, Bensimon G, Meininger V, Leigh P, Guillet P. Dose-ranging study of riluzole in amyotrophic lateral sclerosis. *Lancet*. 1996;347(9013):1425-1431. doi:10.1016/S0140-6736(96)91680-3
- Abe K, Aoki M, Tsuji S, et al. Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2017;16(7):505-512. doi:10.1016/S1474-4422(17)30115-1
- Hardiman O, van den Berg LH. Edaravone: a new treatment for ALS on the horizon? *Lancet Neurol*. 2017;16(7):490-491. doi:10.1016/S1474-4422(17)30163-1
- Meininger V, Asselain B, Guillet P, et al. Pentoxifylline in ALS: a double-blind, randomized, multicenter, placebo-controlled trial. *Neurology*. 2006;66(1):88-92. doi:10.1212/01.WNL.0000191326.40772.62
- Gordon PH, Moore DH, Miller RG, et al. Efficacy of minocycline in patients with amyotrophic lateral sclerosis: a phase III randomised trial. *Lancet Neurol*. 2007;6(12):1045-1053. doi:10.1016/S1474-4422(07)70270-3
- Piepers S, Veldink JH, De Jong SW, et al. Randomized sequential trial of valproic acid in amyotrophic lateral sclerosis. *Ann Neurol*. 2009;66(2):227-234. doi:10.1002/ana.21620
- McDermott CJ, Shaw PJ, Cooper CL, et al. Safety and efficacy of diaphragm pacing in patients with respiratory insufficiency due to amyotrophic lateral sclerosis (DiPALS): a multicentre, open-label, randomised controlled trial. *Lancet Neurol*. 2015;14(9):883-892. doi:10.1016/S1474-4422(15)00152-0
- Gonzalez-Bermejo J, Morelot-Panzini C, Tanguy ML, et al. Early diaphragm pacing in patients with amyotrophic lateral sclerosis (RespiStimALS): a randomised controlled triple-blind trial. *Lancet Neurol*. 2016;15(12):1217-1227. doi:10.1016/S1474-4422(16)30233-2
- van Eijk RPA, Nikolakopoulos S, Ferguson TA, Liu D, Eijkemans MJ, van den Berg LH. Increasing the efficiency of clinical trials in neurodegenerative disorders using group sequential trial designs. *J Clin Epidemiol*. 2018;98:80-88. doi:10.1016/j.jclinepi.2018.02.013
- Wong C, Stavrou M, Elliott E, et al. Clinical trials in amyotrophic lateral sclerosis: a systematic review and perspective. *Brain Commun*. 2021;3(4):fcab242. doi:10.1093/BRAINCOMMS/FCAB242
- Meurer WJ, Tolles J. Interim analyses during group sequential clinical trials. *JAMA*. 2021;326(15):1524-1525. doi:10.1001/JAMA.2021.10174
- Friedman LM, Furberg CD, DeMets D, et al. *Fundamentals of Clinical Trials*. 5th ed. Springer; 2015.
- Nikolakopoulos S, Roes KC, van der Tweel I. Sequential designs with small samples: evaluation and recommendations for normal responses. *Stat Methods Med Res*. 2018;27(4):1115-1127. doi:10.1177/0962280216653778
- Pocock SJ. Group sequential methods in the design and analysis of clinical trials. *Biometrika*. 1977;64(2):191-199. doi:10.2307/2335684
- O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979;35(3):549-556. doi:10.2307/2530245
- O'Hagan A, Stevens JW, Campbell MJ. Assurance in clinical trial design. *Pharm Stat*. 2005;4(3):187-201. doi:10.1002/pst.175
- Tornese P, Lalli S, Cocco A, Albanese A. Review of disease-modifying drug trials in amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. 2022;93(5):521-529. doi:10.1136/jnnp-2021-328470
- Cudkowicz ME, Titus S, Kearney M, et al. Safety and efficacy of ceftriaxone for amyotrophic lateral sclerosis: a multi-stage, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2014;13(11):1083-1091. doi:10.1016/S1474-4422(14)70222-4
- US Food and Drug Administration. Summary of safety and probable benefit (SSPB). 2011. Accessed April 29, 2021. accessdata.fda.gov/cdrh_docs/pdf/10/H100006b.pdf.
- Van Eijk RPA, Nikolakopoulos S, Roes KCB, et al. Critical design considerations for time-to-event endpoints in amyotrophic lateral sclerosis clinical trials. *J Neurol Neurosurg Psychiatry*. 2019;90(12):1331-1337. doi:10.1136/jnnp-2019-320998
- Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Ltd; 2004.
- Berry JD, Sheffer JM, Conwit R, et al. Design and initial results of a multi-phase randomized trial of ceftriaxone in amyotrophic lateral sclerosis. *PLoS One*. 2013;8(4):e61177. doi:10.1371/JOURNAL.PONE.0061177
- 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(2):156-161. doi:10.1136/JNNP-2017-317077
- van Eijk RPA, Nikolakopoulos S, Roes KCB, et al. Innovating clinical trials for amyotrophic lateral sclerosis: challenging the established order. *Neurology*. 2021;97(11):528-536. doi:10.1212/WNL.00000000000012545
- European Medicines Agency. Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design. 2007. Accessed April 20, 2021. ema.europa.eu/en/documents/scientific-guideline/reflection-paper-methodological-issues-confirmatory-clinical-trials-planned-adaptive-design_en.pdf.
- US Food and Drug Administration. *Adaptive Design Clinical Trials for Drugs and Biologics Guidance for Industry*. 2019. Accessed May 6, 2021. fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics-guidance-industry.
- Pocock SJ. When to stop a clinical trial. *BMJ*. 1992;305(6847):235-240. doi:10.1136/BMJ.305.6847.235
- Lachin JM, Foulkes MA. Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics*. 1986;42(3):507-519. doi:10.2307/2531201
- Stevely A, Dimairo M, Todd S, et al. An investigation of the shortcomings of the CONSORT 2010 statement for the reporting of group sequential randomised controlled trials: a methodological systematic review. *PLoS One*. 2015;10(11):e0141104. doi:10.1371/journal.pone.0141104

33. Jennison C, Turnbull BW. *Group Sequential Methods with Applications to Clinical Trials*. Chapman and Hall/CRC; 1999.
34. I AM ALS Clinical Trials Team. *Patient-Centric Trial Design (PaCTD) Rating Criteria*. Accessed April 15, 2022. iamals.org/patient-centric-trial-design-pactd-rating-criteria/.
35. Van Den Berg LH, Sorenson E, Gronseth G, et al. Revised Airline House consensus guidelines for design and implementation of ALS clinical trials. *Neurology*. 2019;92(14):e1610-e1623. doi:10.1212/WNL.00000000000007242
36. Dimairo M, Julious SA, Todd S, Nicholl JP, Boote J. Cross-sector surveys assessing perceptions of key stakeholders towards barriers, concerns and facilitators to the appropriate use of adaptive designs in confirmatory trials. *Trials*. 2015;16(1):585. doi:10.1186/s13063-015-1119-x
37. Chen TT. Statistical issues and challenges in immuno-oncology. *J Immunother Cancer*. 2013;1:18. doi:10.1186/2051-1426-1-18
38. Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N Engl J Med*. 2016;375(1):65-74. doi:10.1056/NEJMRA1510061
39. Burnett T, Mozgunov P, Pallmann P, Villar SS, Wheeler GM, Jaki T. Adding flexibility to clinical trial designs: an example-based guide to the practical use of adaptive designs. *BMC Med*. 2020;18(1):352. doi:10.1186/S12916-020-01808-2
40. van Es MA, Hardiman O, Chio A, et al. Amyotrophic lateral sclerosis. *Lancet*. 2017;390(10107):2084-2098. doi:10.1016/S0140-6736(17)31287-4
41. Psioda MA, Ibrahim JG. Bayesian clinical trial design using historical data that inform the treatment effect. *Biostatistics*. 2019;20(3):400-415. doi:10.1093/biostatistics/kxy009
42. Wu X, Xu Y, Carlin BP. Optimizing interim analysis timing for Bayesian adaptive commensurate designs. *Stat Med*. 2020;39(4):424-437. doi:10.1002/SIM.8414
43. Westenberg 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(5):423-433. doi:10.1016/S1474-4422(18)30089-9
44. van Eijk RPA, Westenberg HJ, Nikolakopoulos S, et al. Refining eligibility criteria for amyotrophic lateral sclerosis clinical trials. *Neurology*. 2019;92(5):e451-e460. doi:10.1212/WNL.0000000000006855
45. Paganoni S, Macklin EA, Hendrix S, et al. Trial of sodium phenylbutyrate-taurursodiol for amyotrophic lateral sclerosis. *N Engl J Med*. 2020;383(10):919-930. doi:10.1056/nejmoa1916945
46. Oki R, Izumi Y, Fujita K, et al. Efficacy and safety of ultrahigh-dose methylcobalamin in early-stage amyotrophic lateral sclerosis: a randomized clinical trial. *JAMA Neurol*. 2022;79(6):575-583. doi:10.1001/JAMANEUROL.2022.0901
47. Wason JMS, Brocklehurst P, Yap C. When to keep it simple: adaptive designs are not always useful. *BMC Med*. 2019;17(1):152. doi:10.1186/s12916-019-1391-9
48. Baxi EG, Thompson T, Li J, et al. Answer ALS, a large-scale resource for sporadic and familial ALS combining clinical and multi-omics data from induced pluripotent cell lines. *Nat Neurosci*. 2022;25(2):226-237. doi:10.1038/s41593-021-01006-0
49. Atassi N, Berry J, Shui A, et al. The PRO-ACT database: design, initial analyses, and predictive features. *Neurology*. 2014;83(19):1719-1725. doi:10.1212/WNL.0000000000000951

Neurology®

Development and Evaluation of a Simulation-Based Algorithm to Optimize the Planning of Interim Analyses for Clinical Trials in ALS

Jordi W.J. van Unnik, Stavros Nikolakopoulos, Marinus J.C. Eijkemans, et al.
Neurology 2023;100:e2398-e2408 Published Online before print April 21, 2023
DOI 10.1212/WNL.0000000000207306

This information is current as of April 21, 2023

Updated Information & Services	including high resolution figures, can be found at: http://n.neurology.org/content/100/23/e2398.full
References	This article cites 42 articles, 10 of which you can access for free at: http://n.neurology.org/content/100/23/e2398.full#ref-list-1
Subspecialty Collections	This article, along with others on similar topics, appears in the following collection(s): Amyotrophic lateral sclerosis http://n.neurology.org/cgi/collection/amyotrophic_lateral_sclerosis Clinical trials Methodology/study design http://n.neurology.org/cgi/collection/clinical_trials_methodology_study_design
Permissions & Licensing	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: http://www.neurology.org/about/about_the_journal#permissions
Reprints	Information about ordering reprints can be found online: http://n.neurology.org/subscribers/advertise

Neurology® is the official journal of the American Academy of Neurology. Published continuously since 1951, it is now a weekly with 48 issues per year. Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology. All rights reserved. Print ISSN: 0028-3878. Online ISSN: 1526-632X.

