

Functional Loss and Mortality in Randomized Clinical Trials for Amyotrophic Lateral Sclerosis: To Combine, or Not to Combine—That is the Estimand

Ruben P.A. van Eijk^{1,2,*}, Kit C.B. Roes³, Inez de Greef-van der Sandt⁴, Leonard H. van den Berg² and Ying Lu¹

Amyotrophic lateral sclerosis is a rapidly progressive disease leading to death in, on average, 3–5 years after first symptom onset. Consequently, there are frequently a non-negligible number of patients who die during the course of a clinical trial. This introduces bias in end points such as daily functioning, muscle strength, and quality of life. In this paper, we outline how the choice of strategy to handle death affects the interpretation of the trial results. We provide a general overview of the considerations, positioned in the estimand framework, and discuss the possibility that not every strategy provides a clinically relevant answer in each setting. The relevance of a strategy changes as a function of the intended trial duration, hypothesized treatment effect, and population included. It is important to consider this trade-off at the design stage of a clinical trial, as this will clarify the exact research question that is being answered, and better guide the planning, design, and analysis of the study.

Amyotrophic lateral sclerosis (ALS) is a relatively rare and rapidly progressive disease causing loss of voluntary muscle function and leading to death in, on average, 3–5 years after first symptom onset.^{1,2} A key challenge in clinical trials for ALS is the extensive clinical heterogeneity between patients,^{3,4} where survival time may range from a few months to over 20 years.⁵ Traditionally, increasing the patient's life expectancy has been the primary objective of randomized clinical trials in ALS. Such studies require, however, relatively long study durations and large sample sizes.^{6,7} In more recent clinical trials, supported by the patient community,⁸ a shift can be observed toward studies with a shorter follow-up.⁹ In these settings, where study durations commonly range from 6 to 12 months, showing an improvement in overall survival has become impractical.¹⁰ This has changed the primary focus of ALS clinical trials to intermediate end points such as daily functioning, muscle strength, or quality of life.

Nevertheless, even in these shorter clinical trials, frequently a non-negligible number of patients die during the course of the study.^{11,12} Although randomization protects the treatment comparison before or at randomization, the occurrence of death after randomization is directly related to the patient's disease progression^{5,13,14} and may be influenced by the treatment itself. As such, these associations with death could introduce bias in study end points that are based on function. At a patient level, we are often left wondering what would have happened had the patient not died: Would ALS have progressed functionally; would the treatment have protected the

patient from further progression; or would death have occurred had the patient been randomized to a different arm.

In the past, there have been extensive discussions about how to handle death in common end points such as the Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised (ALSFRS-R) or vital capacity.^{12,13} Proposed solutions range from simple imputations to complex algorithms which simultaneously integrate death with the end point of interest. An often overlooked aspect, however, is that the choice of how we address death alters the research question and impacts the clinical interpretation of the study results. Therefore, prior to the discussion about which strategy for handling death may be “best” or “most optimal,” there is a need to define the exact research question that we seek to answer.^{15–17}

The question to be answered may not always be straightforward.¹⁷ A patient, for example, may be most interested in how treatment improves their daily functioning during life. The strategy to address death, therefore, must reflect a method that provides the “unbiased” effect of treatment on the patient's daily functioning during life. In contrast, a physician, payer, or regulator may be much more interested in the totality of the treatment effect. As such, a preferred strategy would be a method that summarizes the total treatment effect on daily functioning and survival, for example, by using a composite end point. Both strategies answer a fundamentally different research question, namely: “What is the effect of treatment on daily functioning *during life*?” vs. “What is the effect of treatment on daily functioning *and survival*?” The

¹Department of Biomedical Data Science and Center for Innovative Study Design, School of Medicine, Stanford University, Stanford, California, USA; ²Department of Neurology, UMC Utrecht Brain Centre, University Medical Centre Utrecht, Utrecht, The Netherlands; ³Department of Health Evidence, Section Biostatistics, Radboud Medical Centre Nijmegen, The Netherlands; ⁴Treeway BV, Tilburg, The Netherlands. *Correspondence: Ruben P.A. van Eijk (r.p.a.vaneijk-2@umcutrecht.nl)

Received December 30, 2021; accepted January 15, 2022. doi:10.1002/cpt.2533

answer to each of these questions may be different, possibly leading to different conclusions from the same study results. Clarifying the exact scientific objective of a study will not only help guide the discussion about which strategy is required to handle death, but also direct the planning, design, and analysis of clinical trials, and the definition of primary and secondary objectives.

In this paper, therefore, we outline how our choice to handle death in the analysis may change the interpretation of the trial results, illustrated by clinical trials in ALS. We aim to provide a general overview of the main considerations together with a discussion of the most commonly used strategies. We will position these discussions in the estimand framework.^{15,16,18}

THE ESTIMAND FRAMEWORK

The estimand framework, as outlined in the ICH E9(R1) addendum (International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use Addendum on Estimands and Sensitivity Analysis in Clinical Trials to The Guideline on Statistical Principles for Clinical Trials, E9(R1)),¹⁹ aims to help investigators define the exact research question and to align the trial objectives with the study design, including end points and analysis.^{20,21} The estimand, meaning “what needs to be estimated to address the research question,” can be conceptualized as a systematic framework to phrase the different components of the research question. The estimand includes five attributes: (i) the treatment, i.e., a description of the intervention and its comparator, (ii) the population, i.e., the patients to whom the results should apply, (iii) the variable, i.e., the efficacy end point that is required to answer the research question, (iv) the strategy, i.e., a description of how to account for postrandomization events, and (v) the population-level summary, i.e., the summary statistic that reflects the treatment effect in the efficacy end point.²² An explicit term in the estimand framework is postrandomization event or “intercurrent event.” These are events that occur after treatment initiation, and affect either the interpretation or the existence of the measurements associated with the clinical question of interest.^{18,23} Examples include use of concomitant treatments, changes in background treatments, treatment discontinuation, or the occurrence of terminal events, such as death.

The analytical strategy to account for these intercurrent events directly impacts the estimated treatment effect. Thus, based on the research question that we want to answer, we need to consider what strategy best aligns with our objectives. In general, the estimand framework distinguishes five different (example) strategies to address intercurrent events: (i) treatment policy, (ii) hypothetical, (iii) composite, (iv) while on treatment (hereafter referred to as while alive), and (v) principal stratum.¹⁹ In the treatment policy strategy, one ignores the event and uses the data as is. An example is treatment discontinuation: We would use all available follow-up data of that patient, irrespective of whether the patient actually underwent treatment or not (i.e., similar to the “intention-to-treat” principle). Strictly speaking, we should include the data on the variable of interest after the event. In case of death, however, such a strategy becomes infeasible as the data simply do not exist and we can’t ignore its occurrence.^{20,24,25} Accounting for death using the principal stratum strategy, thereby targeting the patients in whom

death does not occur and estimating counterfactual outcomes between treatment arms, is hardly ever used.²⁶ In this paper, therefore, we will focus primarily on the following strategies to account for death in ALS clinical trials: (i) the hypothetical strategy, envisaging a hypothetical world where death does not occur, (ii) the while-alive strategy, summarizing the treatment effect while the patient is alive, and (iii) the composite strategy, making death part of the primary end point of the study.

THE HYPOTHETICAL STRATEGY

In **Table 1**, we provide an overview of all randomized controlled clinical trials in ALS published between 2018 and 2021, together with the strategy applied to account for death when evaluating the ALSFRS-R, the most commonly used efficacy end point in ALS.⁹ In 6 of the 14 studies (43%), a hypothetical estimand was targeted. The treatment effect being estimated is the expected improvement in ALSFRS-R at the end of the trial, if the patient survives the treatment period. If death rates are low, this estimand has clinical value from a patient perspective, e.g., “If you continue to take the treatment and survive the coming X months, you can expect an efficacy response of Y points in your ALSFRS-R total score at month X.” In this case, the population-level summary is the mean difference in ALSFRS-R total score between treatment and placebo at month X. In order to estimate this effect size, however, all patients must complete the trial. The key challenge for this estimand is, therefore, that the occurrence of death is seen as a missing data problem, and we need to make an assumption for patients who died prior to the end of the study. Importantly, we must assume that the time of death is independent of treatment in order for the estimand to be clinically relevant. In fact, the relevance of this estimand may be the real “hypothetical” part: Can we act hypothetically as if death did not occur and still answer a clinically relevant research question?

As can be seen in **Table 1**, a variety of suggestions has been proposed to impute patient outcomes after death, all leading to potentially different estimates of the treatment effect. In the most simplistic case, we could simply ignore the data of deceased patients and solely use the data from patients who survive until month X (“survivor” or “complete case” analysis). For the population-level summary to be valid, however, we have to assume that the patients who die are not different from patients who survive, and that outcomes would have been similar at the end of the trial. Apart from the fact that this assumption could be contested in ALS, another disadvantage is that one uses only the data from survivors, which leads to loss of statistical power in an already rare disorder. As an alternative, we could attempt to impute the missing observations at month X for deceased patients, where two common methods include (i) last observation carried forward, or (ii) worst score imputation. In the first strategy, we assume that if the patient hadn’t died during our study, the patient’s condition would not have progressed after their last follow-up visit. In the second strategy we assume that all deceased patients would have had the worst ALSFRS-R score observed among surviving patients at month X.

The disadvantage of the above-mentioned methods is that one either makes a group-level assumption, e.g., all patients who die

Table 1 Overview of strategies to handle death in randomized clinical trials in ALS between 2018 and 2021

Author (Year)	Phase	Drug	Sample size	No. deaths (%)	Duration (weeks)	Method	Strategy
Ahmadi (2018)	II	Nanocurcumin	54	5 (9%)	52	Last score	Hypothetical
Aizawa (2021)	II	Perampanel	65	1 (2%)	48	MMRM	Hypothetical
Benatar (2018)	II	Arimoclomol	38	13 (34%)	52	Mixed model	While alive
Berry (2019)	II	Mesenchymal stem cells	48	0 (0%)	24	—	—
Chen (2020)	II	Tamoxifen	18	2 (11%)	52	MMRM	Hypothetical
Cudkowicz (2021)	III	Levosimendan	496	44 (9%)	48	Joint rank	Composite
de la Rubia (2019)	II	EH301	32	2 (6%)	16	Survivor analysis	Hypothetical
Kaji (2019)	III	Methylcobalamin	373	73 (20%)	182	Worst score	Hypothetical
Ludolph (2018)	III	Rasagiline	252	75 (30%)	78	Linear regression	While alive
Meininger (2017)	II	Ozanezumab	303	14 (5%)	48	Joint rank	Composite
Mora (2019)	III	Masitinib	394	33 (8%)	48	Zero score	Composite
Paganoni (2020)	II	Phenylbutyrate-Taurursodiol	137	7 (5%)	24	Mixed model	While alive
Statland (2019)	II	Rasagiline	80	9 (11%)	52	Mixed model	While alive
van Es (2020)	II	Penicillin G-Hydrocortisone	16	6 (38%)	52	Joint model	While alive
Vucic (2021)	II	Dimethyl fumarate	107	1 (1%)	36	Multiple imputation	Hypothetical

Randomized, placebo-controlled clinical trials, published between January 2018 and November 2021, which included at least 12 weeks of treatment. Crossover studies were excluded.

ALS, amyotrophic lateral sclerosis; MMRM, mixed model for repeated measures.

receive score Z at month X , or that we make a too simplistic individual assumption in a known progressive and highly heterogeneous disorder. As we will discuss next in “the while-alive strategy,” mixed models for repeated measures (MMRM) partially overcome these limitations. The MMRM assumes that a patient who dies would have had similar outcomes at month X compared with a patient with similar ALSFRS-R scores up to the time of death, but who survived.²⁷

THE WHILE-ALIVE STRATEGY

The critical issue with the hypothetical estimand is the fact that it loses its relevance when death rates increase, irrespective of which method is chosen. For example, if only 50% of the patients survive until month X , how relevant is it for a patient to know their expected improvement in ALSFRS-R at month X if they are unlikely to be alive at that time anyway? Therefore, as an alternative, we could estimate the treatment effect while the patient is alive. This changes the information communicated to the patient. For example: “If you continue to take the treatment while you are alive, you can expect an efficacy response of $Y\%$ reduction in your ALSFRS-R progression rate after X months or until death (whichever occurs first).” In this case, the population-level summary has changed from a mean difference at month X in the hypothetical estimand to a percent reduction in progression rate after X months or until death (whichever occurs first). Note that we also switch in the while-alive strategy to progression rates rather than the actual value of our end point, and the effect size is no longer linked to any particular timepoint. As a

consequence, the treatment effect applies to a patient who survives X months, but also to a patient who dies prior to that time.

From **Table 1** we can observe that the while-alive estimand was targeted in 36% of the ALS clinical trials. Regression models are a helpful method for estimating the average progression rate in the treatment and placebo arm. The treatment effect, subsequently, can be simply summarized as either an absolute or relative reduction in the disease progression rate compared with placebo. In general, there are currently two main methods for calculating the average progression rate: a two-stage approach, calculating, for each patient, their individual progression rate and averaging rates across patients (e.g., Ludolph, 2018, **Table 1**), or a one-stage approach, defining a random-slopes model with random effects per patient (e.g., Benatar, 2018, **Table 1**).

The difference between the two methods is illustrated in **Figure 1** for a hypothetical patient. The disadvantage of the two-stage approach is that when only a few data points are available, this could lead to extreme estimates (e.g., **Figure 1d**). Moreover, while the patient who dies will provide far less information overall, when estimating the group average rate of decline, a deceased patient and a patient who survives contribute equally. The one-stage approach using a random-slopes model improves on this by adjusting, or shrinking, the estimated patient-specific regression line toward the population mean, based on the amount of information available. The MMRM is similar to the random-slopes model, with the main difference being an MMRM models time as categorical variable. As a consequence, the MMRM provides

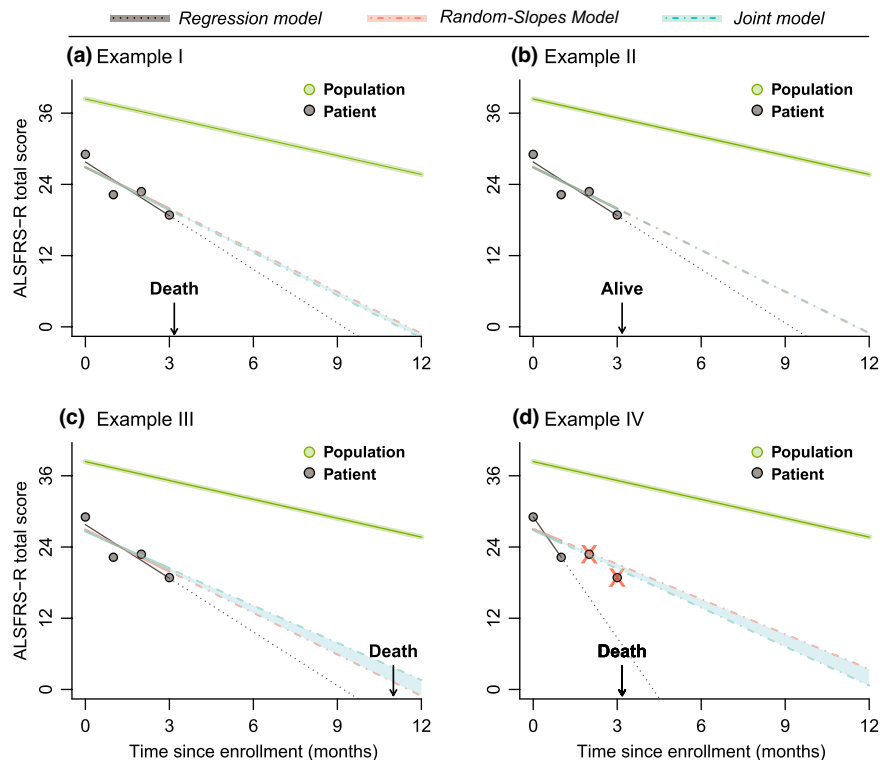


Figure 1 Comparison of different models for the estimated patient trajectory after death. Illustration of hypothetical patient data with varying reasons for and timing of dropout. The green line represents the overall population trajectory over time (e.g., the average of all patients in a well-defined cohort, or patients with ALS allocated to a treatment arm). (a) In the first scenario, the patient dies between Month 3 and Month 4. As can be seen, in a random-slopes model (red), the patient’s individual regression line (black) is drawn or shrunk toward the population average, based on the amount of information available. (a–c) A joint model (blue) provides a similar estimate compared with the random-slopes model, but alters the shrinkage factor based on the reason and timing of dropout. The blue shaded area highlights the difference between the random-slopes and joint model. In scenario (d) we illustrate how the model estimates change if the last two observations prior to death are not observed. ALSFRS-R, Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

the mean difference at a certain timepoint or visit, reflecting the hypothetical estimand, whereas a random-slopes model provides the difference in progression rates over a certain time period. Hence, the main benefit of the random-slopes model is that it provides an effect estimate applicable not only to patients who die prior to a certain timepoint but also to those who survive. In contrast, the MMRM provides an effect estimate applicable only to those patients who survive until a certain timepoint. A disadvantage of the random-slopes model is, however, that we not only need to make an assumption about the time trajectory (e.g., linear, quadratic, cubic) but also about the treatment effect over time, which is not required for the MMRM.

In **Figure 1** we can see that the reason why data are missing, or the timing of death occurring do not affect the one-stage or two-stage approach. For the model, it simply does not matter whether the patient dies after the last observed data point, or whether they simply withdraw from the study. This may introduce bias,^{28,29} and—also from a clinical perspective—it would be important to consider why the data are missing. This additional information can be incorporated by using a different class of models called shared-parameter or joint models.^{12,13,30} A joint model is identical to a random-slopes model, but rather than optimizing model parameters using only the ALSFRS-R data, it also incorporates information on survival time. As such, model

parameters, including the individual trajectory of the patient, are estimated such that the model optimally represents both the ALSFRS-R and survival data.^{30,31} Due to this joint optimization process, the timing of death, or the reason for dropout, lead to different patient trajectories and may prevent bias when used appropriately.^{12,29}

A SIMULATED COMPARISON

By means of simulation, we will illustrate the impact of each of the discussed methods in a hypothetical clinical trial setting. In this setting, treatment has no effect on the ALSFRS-R but produces harmful side-effects, doubling the hazard for death in the treatment arm. This results in a 12-month survival probability of 85% for patients allocated to placebo, and 75% for patients allocated to treatment. Suppose we enroll 100 patients per arm, where ALSFRS-R is collected at months 0, 1, 2, 4, 6, 8, 10, and 12. Furthermore, let us assume that, based on previous literature,¹³ patients decline on average by 1.06 ALSFRS-R points per month, where each point loss in ALSFRS-R increases the immediate risk of death by 13%. Based on these settings, the change from baseline is, on average, 12.72 ALSFRS-R points for both treatment arms. **Table 2** provides the estimated mean per treatment arm, together with the treatment effect, for each method. For the sake of illustration, we made the while-alive estimand comparable to

Table 2 Estimated treatment effect on functional decline for a hypothetical randomized controlled trial with fatal side effects

Estimand	Mean placebo (at month 12)	Mean treatment (at month 12)	Mean difference (treated minus placebo)	One-sided rejection proportion	
				Difference ≤ 0	Difference ≥ 0
Truth (generated via joint model)	-12.72	-12.72	0.000	0.0250	0.0250
Hypothetical estimand					
Survivor analysis (complete case)	-11.47	-10.64	0.838	0.0058	0.0771
Last observation carried forward	-11.48	-10.58	0.900	0.0037	0.1028
Worst score imputation	-15.01	-16.82	-1.813	0.1328	0.0024
Mixed model for repeated measures	-11.56	-11.45	0.104	0.0193	0.0291
While-alive estimand					
Regression model (two-stage)	-12.85	-12.92	-0.070	0.0235	0.0256
Random-slopes model (one-stage)	-12.66	-12.59	0.067	0.0219	0.0290

The while-alive estimand targeted the progression rate per month; results were multiplied by 12 to make them comparable. Results are based on 100,000 iterations; the joint modeling framework was used to simulate conditional ALSFRS-R and survival data; treatment resulted in a hazard ratio of 2.0; simulation parameters are described elsewhere (100 patients per arm).¹³

ALSFRS-R, Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

the hypothetical estimand by multiplying the estimated monthly progression rate by 12.

For the hypothetical estimand, the survivor analysis, last observation carried forward, and worst case imputation led to considerable deviating estimates, falsely suggesting a beneficial or harmful effect on ALSFRS-R. The bias was reduced when using the MMRM, but underestimation of the mean change from baseline remains. To put the observed differences into context: a 2.5-point difference after 24 weeks was considered to be sufficient to market a therapy for ALS in several countries.³² The one-stage and two-stage approaches resulted, in this simulation setting, in a closer approximation to the truth, though results may change when simulating different missing data mechanisms or nonlinear functional trajectories.^{33,34} Interestingly, the MMRM and random-slopes model rejected the null hypothesis of no effect slightly more often in favor of the treatment even though treatment has no effect on function and shortens survival time. Though the differences are small, these findings are comparable to a previous study¹² and demonstrate the potential problems when our assumptions are not entirely accurate.

THE COMPOSITE STRATEGY

As an alternative to the while-alive and hypothetical strategy, we can integrate death with the variable of interest and make death part of the primary end point definition. Consequently, the composite analytical strategy does not require data after death.³⁵ An important objective is to define a composite with the ALSFRS-R in which the occurrence of death does not result in a “good” outcome.^{24,25} A simple method, for example, is to score deceased patients worse than patients who survive prior to ranking their ALSFRS-R outcomes. As a consequence, however, the treatment effect that we estimate in a composite estimand no longer solely reflects the improvement in function but also contains the improvement in survival.

In ALS clinical trials, two common composite strategies include (i) giving all deceased patients the worst theoretical score

on ALSFRS-R (0), or (ii) ranking patients based on their time of death and ALSFRS-R outcome. The first strategy scores all deceased patients equally and does not distinguish when the patient has died during the study. Rank-based methods, such as the combined assessment of function and survival (CAFS),¹¹ or variants thereof,³⁶ improve on this by giving the first patient who dies the lowest score, and the patient who survives with the best ALSFRS-R outcomes the highest score. This reduces the number of “tied” observations and increases statistical power.

Nevertheless, a disadvantage of rank-based methods, and composite strategies in general, is that the interpretation and quantification of the treatment effect is less straightforward.³⁵ For a rank test, we could express the effect size as a winning probability,³⁷ reflecting the probability that a random patient in the treated arm has a better outcome than a random patient in the placebo arm. This probability, however, may be difficult to grasp for a patient, especially since the effect could be driven by either an improvement in function or survival, and it does not provide an exact quantity that is relatable to daily life. The time to a composite event end point, such as the time to a 6-point decrease in ALSFRS-R or death,³⁸ could be considered as an alternative, and may be easier to interpret, as it reflects the probability of an unfavorable event in the following X months, regardless of survival.²⁴ A composite event, however, considers each component to be of equal importance, which may not always be appropriate when the clinical impact of components differs significantly.

Importantly, most of the composite methods use only part of the data: Patient outcomes are based on either their survival time *or* their ALSFRS-R score. This may not only lead to a suboptimal use of information, but could also “dilute” treatment effects if the components of the composite are contributing disproportionately. For example, if 20% of the patients die at the end of the trial, 80% of the ranking scores are based on ALSFRS-R data. If treatment selectively improves survival and not the ALSFRS-R, the treatment effect that we observe on the composite end point can only

Table 3 Summary of the hypothetical, while-alive, and composite estimand for ALS clinical trials

	Hypothetical strategy	While-alive strategy	Composite strategy
Example re-search question	In patients with ALS who survive 12 months, what is the between-group difference in mean daily functioning, as measured by the ALSFRS-R total score, between treatment and placebo, at 12 months after randomization?	In patients with ALS, what is the between-group difference in mean rate of functional loss, as measured by the ALSFRS-R total score, between treatment and placebo, over 12 months after randomization or until death (whichever occurs first)?	In patients with ALS, what is the probability that a random patient on treatment has a longer survival or better daily functioning, as measured by the ALSFRS-R total score, compared with a random patient on placebo, at 12 months of treatment?
Benefits	<ul style="list-style-type: none"> • Clinically relevant and easy to explain to patient if survival probability is high • Powerful strategy when death rates are low 	<ul style="list-style-type: none"> • Clinically relevant and easy to explain to patient, irrespective of survival probability 	<ul style="list-style-type: none"> • Depending on method, fewer assumptions and no need to extrapolate after death • Estimates the totality of the treatment effect on both ALSFRS-R and survival
Disadvantages	<ul style="list-style-type: none"> • Assumptions are needed about the patient-specific trajectory before and after death, which can introduce bias in effect estimate • Interpretation of and applicability to the intended population becomes complicated when death rates are high 	<ul style="list-style-type: none"> • Assumptions are required about the patient-specific trajectory over time • Assumptions are required about the response to treatment over time 	<ul style="list-style-type: none"> • Interpretation of the treatment effect is less straightforward and can be driven by each component • Loss of information and/or power when death rates are low, or treatment affects only one component

ALS, amyotrophic lateral sclerosis; ALSFRS-R, Amyotrophic Lateral Sclerosis Functional Rating Scale–Revised.

be driven by 20% of the patients, while the remaining 80% of the patients provide no meaningful information for the effect estimate. Multivariate methods, such as a joint model that incorporates all available information for ALSFRS-R and survival, may provide a more powerful alternative,¹³ but this may be at the price of increasing the risk of defining an erroneous model.³⁹

TO COMBINE, OR NOT TO COMBINE?

Overall, each of the three strategies to address death have strengths and weaknesses, irrespective of which method is chosen (Table 3). For clinical trials with a longer follow-up duration, where death rates are expected to exceed 20–30%,¹⁰ the hypothetical estimand may be the least informative for patients and physicians as it reflects an imaginary world applicable to only a few patients. Nevertheless, the hypothetical estimand holds value for shorter studies with a negligible number of deaths as it may require fewer assumptions about the treatment effect than the while-alive estimand, and its interpretation is more straightforward than for a composite. The primary choice for longer, pivotal clinical trials in ALS seems, however, to be composed of either the while-alive strategy or a strategy that combines the ALSFRS-R and mortality into a composite end point.

An argument in favor of the composite strategy is that it reflects the totality of the treatment effect. Especially in a disease such as ALS, for which there is an urgent and enormous unmet medical need, one may prioritize simple “yes/no” decisions about the success of the new treatment over the exact interpretation of its benefits.¹⁵ Suppose a treatment reduces the ALSFRS-R progression rate but at the same time affects an important prognostic mechanism that is not captured by the ALSFRS-R (e.g., weight loss or cognitive decline). Solely measuring the ALSFRS-R would only partially capture the treatment response and may falsely discard an efficacious drug. In these settings, capturing the totality of the treatment

effect on a composite of ALSFRS-R and survival holds clear advantages over the while-alive strategy.

The success of the composite strategy, however, depends on how well one is able to capture the treatment effects on the individual components. If one of the components is nearly unquantifiable, the effect observed on the composite end point can only be driven by other components. The same holds true for the ALSFRS-R and survival: In a 6-month study on 100 patients, the expected number of deaths is 0 to 5.¹⁰ As such, any treatment effect on the composite is virtually solely driven by the ALSFRS-R. Under these circumstances, even if treatment improves additional prognostic mechanisms, the disadvantages of using a composite may not outweigh the benefits of other strategies.

Thus, the choice to combine the ALSFRS-R and survival is a trade-off between the disadvantages and advantages of each strategy. This is for an important part driven by the number of deaths and depends, therefore, on the underlying survival distribution and the duration of the study.¹⁰ This trade-off is illustrated in Figure 2 for a while-alive strategy using a random-slopes model, and a composite strategy using the CAFS in six hypothetical scenarios with different survival rates. In each scenario, treatment reduces the ALSFRS-R progression rate by 30%. In scenarios **b**, **d**, and **f**, treatment additionally affects a prognostic mechanism not captured by the ALSFRS-R. As can be seen, even though the CAFS captures the totality of the treatment effect, this only leads to improved statistical power over the while-alive strategy in longer trials, or when death rates are high. On the other hand, if there is no additional prognostic benefit of treatment, or survival rates are low, the CAFS underperforms due to the loss of information.^{11–13} As such, the decision to use a composite strategy, especially when the primary goal is to find an efficacious drug, should be carefully weighed against the hypothesized treatment benefit, intended duration of the trial, and the targeted patient population.

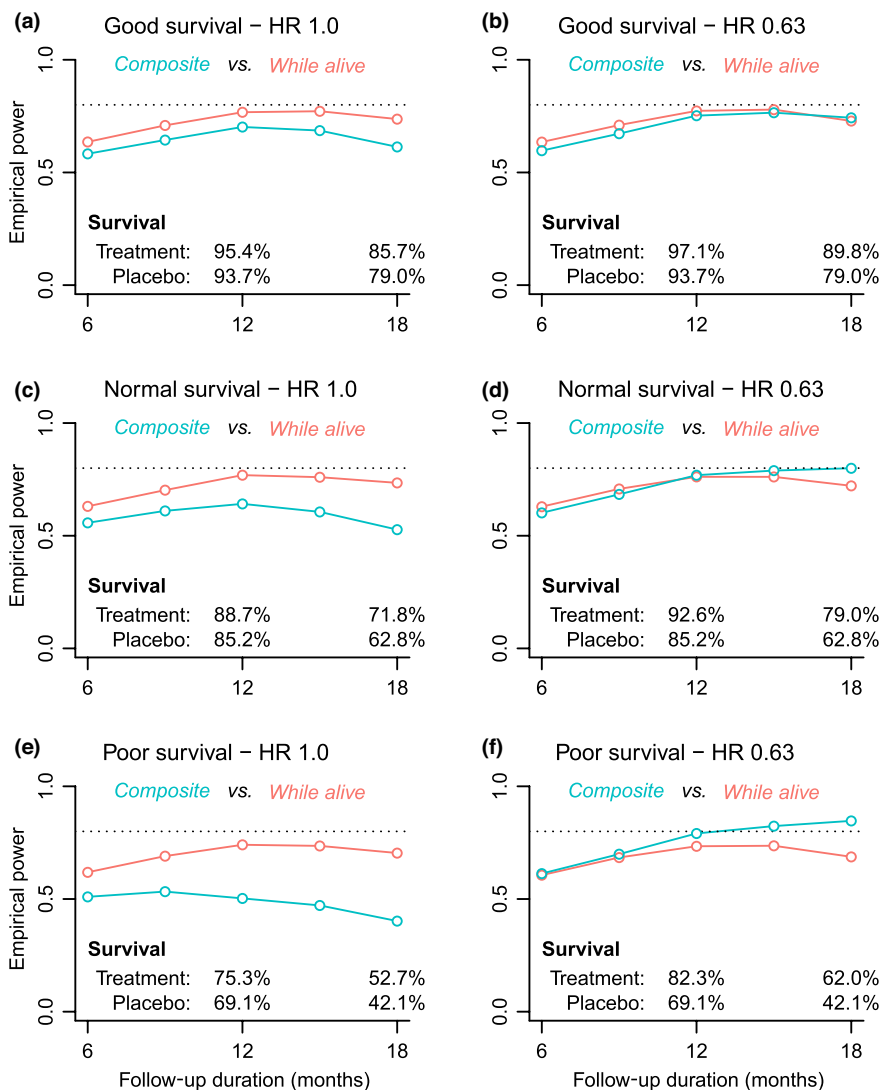


Figure 2 Empirical power of the while-alive and composite strategies. Data were simulated for a hypothetical randomized controlled clinical trial where treatment reduces ALSFRS-R progression rate by 30% (**a, c, and e**). In scenarios (**b, d, and f**), an additional treatment effect on survival, independent of the ALSFRS-R, was added with an HR of 0.63.⁴⁰ The composite strategy was based on the CAFS (combined assessment of function and survival) end point;¹¹ the while-alive strategy was approached using a random-slopes model. Each scenario was simulated 10,000 times; exact simulation details are provided elsewhere.¹³ HR, hazard ratio.

FINAL REMARKS

In this paper, we have provided a general overview of the considerations that play a role when addressing death in efficacy end points for ALS clinical trials. The key consideration is that the chosen strategy has a major impact on the research question being answered. Not every analytical strategy will necessarily provide a clinically relevant answer; its relevance changes as a function of the intended trial duration, hypothesized treatment effect, and included population. Therefore, clarifying at the design stage exactly what the research question is that one aims to answer will guide the discussion about which analytical strategy for handling death is required and will better direct the planning, design, and analysis of the study.

It is important to note that we have only discussed the strategy for addressing death. The estimand for other intercurrent events, e.g., treatment discontinuation due to adverse events or nonadherence, require different strategies. In the end, the final estimand of

a trial will be composed of different strategies to address each intercurrent event individually.^{22,25} Moreover, in this paper we have discussed methods commonly used to address death in ALS clinical trials; we did not aim to provide an exhaustive list of all possible options, or provide an in-depth review of each method. Future work can play a major role in this regard. Innovative methods may overcome the limitations of current analytical strategies and allow further scrutiny of when a particular strategy might be most beneficial. The simulations provided in this paper are intended simply as an illustration, and the trade-off between strategies may change when alternating the expected treatment response, the relationship between function and mortality, and using different design settings.

ACKNOWLEDGMENTS

The authors would like to thank Brenda Vollers-King for assistance with language editing.

FUNDING

R.P.A.v.E. is supported by the Dutch Research Council (Rubicon, 452019301); Y.L. is partially supported by the National Institutes of Health Research (5UL1TR003142-03, 5R01HL08977810, 3P30CA12443512).

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

© 2022 The Authors. *Clinical Pharmacology & Therapeutics* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

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

- van Es, M.A. *et al.* Amyotrophic lateral sclerosis. *Lancet* **390**, 2084–2098 (2017).
- Brown, R.H. & Al-Chalabi, A. Amyotrophic lateral sclerosis. *N. Engl. J. Med.* **377**, 162–172 (2017).
- Mitsumoto, H., Brooks, B.R. & Silani, V. Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol.* **13**, 1127–1138 (2014).
- van den Berg, L.H. *et al.* Revised Airlie House consensus guidelines for design and implementation of ALS clinical trials. *Neurology* **92**, e1610–e1623 (2019).
- Westeneng, H.J. *et al.* Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol.* **17**, 423–433 (2018).
- Rutkove, S.B. Clinical measures of disease progression in amyotrophic lateral sclerosis. *Neurotherapeutics* **12**, 384–393 (2015).
- Goyal, N.A. *et al.* Addressing heterogeneity in amyotrophic lateral sclerosis clinical trials. *Muscle Nerve* **62**, 156–166 (2020).
- Andrews, J.A., Bruijn, L.I. & Shefner, J.M. ALS drug development guidances and trial guidelines: consensus and opportunities for alignment. *Neurology* **93**, 66–71 (2019).
- van Eijk, R.P.A., Kliest, T. & van den Berg, L.H. Current trends in the clinical trial landscape for amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* **33**, 655–661 (2020).
- van Eijk, R.P.A. *et al.* Critical design considerations for time-to-event endpoints in amyotrophic lateral sclerosis clinical trials. *J. Neurol. Neurosurg. Psychiatry* **90**, 1331–1337 (2019).
- Berry, J.D. *et al.* The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. *Amyotroph. Lateral Scler. Frontotemporal. Degener.* **14**, 162–168 (2013).
- Healy, B.C. & Schoenfeld, D. Comparison of analysis approaches for phase III clinical trials in amyotrophic lateral sclerosis. *Muscle Nerve* **46**, 506–511 (2012).
- van Eijk, R.P.A., Eijkemans, M.J.C., Rizopoulos, D., van den Berg, L.H. & Nikolakopoulos, S. Comparing methods to combine functional loss and mortality in clinical trials for amyotrophic lateral sclerosis. *Clin. Epidemiol.* **10**, 333–341 (2018).
- Daghlas, S.A., Govindarajan, R., & Pooled Resource Open-Access, A.L.S. Clinical Trials Consortium. Relative effects of forced vital capacity and ALSFRS-R on survival in ALS. *Muscle Nerve* **64**, 346–351 (2021).
- Ruberg, S.J. & Akacha, M. Considerations for evaluating treatment effects from randomized clinical trials. *Clin. Pharmacol. Ther.* **102**, 917–923 (2017).
- Akacha, M. & Kothny, W. Estimands: a more strategic approach to study design and analysis. *Clin. Pharmacol. Ther.* **102**, 894–896 (2017).
- Akacha, M., Bretz, F. & Ruberg, S. Estimands in clinical trials - broadening the perspective. *Stat. Med.* **36**, 5–19 (2017).
- Pétavy, F., Guizzaro, L., Antunes dos Reis, I., Teerenstra, S. & Roes, K.C.B. Beyond "Intent-to-treat" and "Per protocol": Improving assessment of treatment effects in clinical trials through the specification of an estimand. *Br. J. Clin. Pharmacol.* **86**, 1235–1239 (2020).
- Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials: E9(R1). <https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf> (2019). Accessed November 20, 2021.
- Kahan, B.C. *et al.* Treatment estimands in clinical trials of patients hospitalised for COVID-19: ensuring trials ask the right questions. *BMC Med.* **18**, 286 (2020).
- Alemayehu, D., Hemmings, R., Natarajan, K. & Roychoudhury, S. Perspectives on virtual (Remote) clinical trials as the "New Normal" to accelerate drug development. *Clin. Pharmacol. Ther.* **111**, 373–381 (2022).
- Fiero, M.H. *et al.* Demystifying the estimand framework: a case study using patient-reported outcomes in oncology. *Lancet Oncol.* **21**, e488–e494 (2020).
- Oude Rengerink, K., Mitroiu, M., Teerenstra, S., Pétavy, F. & Roes, K.C.B. Rethinking the intention-to-treat principle: one size does not fit all. *J. Clin. Epidemiol.* **125**, 198–200 (2020).
- Schmidli, H., Roger, J.H. & Akacha, M. Estimands for recurrent event endpoints in the presence of a terminal event. *Stat. Biopharm. Res.* (2021).<https://doi.org/10.1080/19466315.2021.1895883>.
- Lawrance, R. *et al.* What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials? *J. Patient Rep. Outcomes* **4**, 68 (2020).
- Mitroiu, M., Oude Rengerink, K., Teerenstra, S., Pétavy, F. & Roes, K.C.B. A narrative review of estimands in drug development and regulatory evaluation: old wine in new barrels? *Trials* **21**, 671 (2020).
- Ratitch, B. *et al.* Defining efficacy estimands in clinical trials: examples illustrating ICH E9(R1) guidelines. *Ther. Innov. Regul. Sci.* **54**, 370–384 (2020).
- Rouanet, A., Helmer, C., Dartigues, J.-F. & Jacqmin-Gadda, H. Interpretation of mixed models and marginal models with cohort attrition due to death and drop-out. *Stat. Methods Med. Res.* **28**, 343–356 (2019).
- García-Hernandez, A., Pérez, T., Pardo, M.D.C. & Rizopoulos, D. MMRM vs joint modeling of longitudinal responses and time to study drug discontinuation in clinical trials using a "de jure" estimand. *Pharm. Stat.* **19**, 909–927 (2020).
- Rizopoulos, D. *Joint models for longitudinal and time-to-event data: with applications in R* (Chapman and Hall/CRC Press, Boca Raton, 2012).
- Wulfsohn, M.S. & Tsiatis, A.A. A joint model for survival and longitudinal data measured with error. *Biometrics* **53**, 330–339 (1997).
- Writing Group; Edaravone (MCI-186) ALS 19 Study Group. Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* **16**, 505–512 (2017).
- Chen, Y.F., Ni, X., Fleisher, A.S., Zhou, W., Aisen, P. & Mohs, R. A simulation study comparing slope model with mixed-model repeated measure to assess cognitive data in clinical trials of Alzheimer's disease. *Alzheimers Dement. (NY)* **4**, 46–53 (2018).
- Ashbeck, E.L. & Bell, M.L. Single time point comparisons in longitudinal randomized controlled trials: power and bias in the presence of missing data. *BMC Med. Res. Methodol.* **16**, 43 (2016).
- Keene, O.N. Strategies for composite estimands in confirmatory clinical trials: examples from trials in nasal polyps and steroid reduction. *Pharm. Stat.* **18**, 78–84 (2019).
- Meininger, V. *et al.* Safety and efficacy of ozanezumab in patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Neurol.* **16**, 208–216 (2017).
- van Eijk, R.P.A., van den Berg, L.H. & Lu, Y. Composite endpoint for ALS clinical trials based on patient preference: Patient-Ranked Order of Function (PROOF). *J. Neurol. Neurosurg. Psychiatry* (2021).<https://doi.org/10.1136/jnnp-2021-328194>.

38. Aggarwal, S.P. *et al.* Safety and efficacy of lithium in combination with riluzole for treatment of amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* **9**, 481–488 (2010).
39. Thomadakis, C., Meligkotsidou, L., Pantazis, N. & Touloumi, G. Longitudinal and time-to-drop-out joint models can lead to seriously biased estimates when the drop-out mechanism is at random. *Biometrics* **75**, 58–68 (2019).
40. Cudkowicz, M.E. *et al.* Dexamipexole versus placebo for patients with amyotrophic lateral sclerosis (EMPOWER): a randomised, double-blind, phase 3 trial. *Lancet Neurol.* **12**, 1059–1067 (2013).