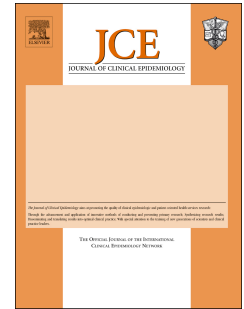


# Journal Pre-proof

Joint modelling of endpoints can be used to answer various research questions in randomized clinical trials

Ruben P.A. van Eijk, Kit C.B. Roes, Leonard H. van den Berg, Ying Lu



PII: S0895-4356(22)00070-1

DOI: <https://doi.org/10.1016/j.jclinepi.2022.03.009>

Reference: JCE 10786

To appear in: *Journal of Clinical Epidemiology*

Received Date: 5 December 2021

Revised Date: 27 February 2022

Accepted Date: 21 March 2022

Please cite this article as: van Eijk RPA, Roes KCB, van den Berg LH, Lu Y, Joint modelling of endpoints can be used to answer various research questions in randomized clinical trials, *Journal of Clinical Epidemiology* (2022), doi: <https://doi.org/10.1016/j.jclinepi.2022.03.009>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 The Author(s). Published by Elsevier Inc.

## **AUTHOR STATEMENT**

**Ruben P.A. van Eijk:** Funding acquisition, Conceptualization, Methodology, Formal analysis, Writing - original draft

**Kit C.B. Roes:** Conceptualization, Methodology, Writing - review & editing

**Leonard H. van den Berg:** Data curation, Resources, Writing - review & editing

**Ying Lu:** Funding acquisition, Conceptualization, Methodology, Supervision, Writing - review & editing

## **Joint modelling of endpoints can be used to answer various research questions in randomized clinical trials**

Ruben P.A. van Eijk (0000-0002-7132-5967),<sup>1,2</sup> Kit C.B. Roes (0000-0002-6775-1963),<sup>3</sup>  
Leonard H. van den Berg (0000-0002-5203-9674),<sup>2</sup> Ying Lu (0000-0002-7698-8962)<sup>1</sup>

(1) Department of Biomedical Data Science and Center for Innovative Study Design, School of Medicine, Stanford University, Stanford, United States.

(2) Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht, the Netherlands.

(3) Department of Health Evidence, Section Biostatistics, Radboud Medical Center Nijmegen, the Netherlands.

### **Corresponding author**

Ruben P.A. van Eijk

Department of Neurology

UMC Utrecht Brain Center, University Medical Center Utrecht

Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands.

Email: [r.p.a.vaneijk-2@umcutrecht.nl](mailto:r.p.a.vaneijk-2@umcutrecht.nl)

Tel: +31 (0) 88 75 554 94; Fax: +31 (0) 88 75 554 94

**Financial support:** RPAvE is supported by the Dutch Research Council (Rubicon, 452019301); YL is partially supported by the National Institutes of Health Research (5UL1TR003142-03, 5R01HL08977810, 3P30CA12443512).

**ABSTRACT**

**Objective:** Correlated longitudinal and time-to-event outcomes, such as the rate of cognitive decline and the onset of Alzheimer's disease, are frequent (co-)primary and key secondary endpoints in randomized clinical trials (RCTs). Despite their biological associations, these types of data are often analyzed separately leading to loss of information and increases in bias. In this paper, we set out how joint modelling of longitudinal and time-to-event endpoints can be used in RCTs to answer various research questions.

**Study Design and Setting:** The key concepts of joint models are introduced and illustrated for a completed trial in amyotrophic lateral sclerosis.

**Results:** The output of a joint model can be used to answer different clinically relevant research questions, where the interpretation of effect estimates and those obtained from conventional methods are similar. Albeit joint models have the potential to overcome the limitations of commonly used alternatives, they require additional assumptions regarding the distributions as well as the associations between two endpoints.

**Conclusion:** Improving the uptake of joint models in RCTs may start by outlining the exact research question one seeks to answer, thereby determining how best to prespecify the model and defining which parameter should be of primary interest.

**KEYWORDS:** Joint modelling, time-to-event, longitudinal, clinical trials, amyotrophic lateral sclerosis, informative censoring.

**RUNNING TITLE:** Joint modelling in randomized clinical trials

## HIGHLIGHTS

- Time-to-event and longitudinal endpoints commonly co-occur in clinical trials
- Joint modelling is a powerful approach to better understand the treatment effect
- Joint models may help to address informative censoring in longitudinal endpoints
- Efficiency gains can be achieved by simultaneous assessment of all available data
- Limitations are increased complexity and the risk of defining an erroneous model

## WHAT IS NEW?

- Joint modelling of time-to-event and longitudinal outcomes have been widely adopted in observational studies, and have been shown to be of benefit in trial settings, but their prospective use in clinical trial practice remains negligible
- The output of a joint model can be used to address different research questions in clinical trials. Connecting joint models with clinically relevant estimands, thereby fine-tuning exactly which clinical question is being answered, could be an important step towards a better understanding of the interaction between treatment and the disease
- Joint models have the potential to overcome the limitations of some commonly used conventional methods in clinical trials and may provide more efficient estimates of the treatment effect

## 1. Introduction

Time-to-event and longitudinal outcomes are common endpoints in randomized clinical trials (RCTs). Clinical events may vary from the onset of a particular disease to the occurrence of disease-related events such as hospitalization, adverse reactions or death. As the time to reach the event may be considerable, it is common practice in RCTs to collect longitudinal information on intermediate outcomes such as physical functioning and quality of life alongside the time-to-event endpoint. In many instances, the longitudinal outcomes themselves are important (co-)primary or key secondary endpoints for evaluating the benefit of treatment. These outcomes are often strongly related to the event of interest; for example, cognitive decline and the onset of Alzheimer's disease [1], or a poorer quality of life and an increased risk of death in oncology trials [2]. Nevertheless, time-to-event and longitudinal endpoints obtained in RCTs are often analyzed separately without consideration of their interrelationships. This may not only lead to a lower consistency and efficiency of estimating treatment effects in RCTs, but also to increased bias and suboptimal use of information [3, 4].

Combined assessment of longitudinal and time-to-event data using a joint model is a powerful approach to better characterize the effect of treatment [3-5]. Joint modelling refers to the simultaneous assessment of two or more outcomes while accounting for their interrelationships using a statistical model [6]. Typically, time-to-event endpoints are assessed using Cox models, whereas linear mixed effects models are commonly used for longitudinal data. In the joint modelling framework, the linear mixed model is linked (or joined) to the risk for the event. Joint models thereby explicitly address the association between the longitudinal and time-to-event endpoint [6]. One can use joint models not only to investigate how a longitudinal variable is related to the event of interest, but also utilize its value in RCTs to address informative missing data in longitudinal endpoints [7, 8], obtain

more efficient estimates of the treatment effect [2, 3, 9], or get a better understanding of the interaction between treatment and the disease [1, 10, 11].

Despite these clear benefits, and the widescale adoption of joint models in observational studies [12, 13], their prospective use in RCTs remains negligible. **Table 1** provides an overview of the statistical analysis conducted in 61 consecutive RCTs published in three major medical journals. Only three RCTs (4.9%) considered a combined analysis of endpoints, though none of them used a joint model. The main barriers for implementation in RCTs may be attributed to perceived difficulties in the interpretation of effect estimates and the apparent complexity of the model [14, 15]. In this paper, therefore, we aim to explain the key concepts of joint models, and set out how joint modelling of longitudinal and time-to-event endpoints can be used in RCTs to answer various research questions, illustrated for Amyotrophic Lateral Sclerosis (ALS).

## 2.1 The valproic acid trial

The valproic acid (VPA) study in ALS will be used to illustrate the joint modelling framework [16]. ALS is a progressive neurological disorder leading to severe muscle weakness and eventually death. Median survival is three to five years, but can range from a few months to over several decades [17]. The extent of physical functioning loss, and its rate of progression, are strongly related to survival time [17, 18]. Physical functioning is commonly quantified by a 12-item questionnaire, the ALS functional rating scale (ALSFRS-R) [19], with lower scores reflecting poorer function. While improving overall survival is the ultimate objective [20-22], preventing functional loss may be of equal importance to patients. Moreover, one could hypothesize that a treatment that reduces or prevents functional decline would also prolong survival time. In **Fig. 1** we provide the observed individual ALSFRS-R

trajectories and survival curves in each treatment arm of the VPA study. The primary objective was to evaluate the effect of VPA on overall survival compared to placebo, with the ALSFRS-R as key secondary endpoint. The trial followed a sequential design; it was stopped early for futility and possible harm due to VPA.

## 2.2 The joint modelling framework

In **Fig. 2A** we provide the ALSFRS-R scores for three placebo-allocated patients in the VPA study. As can be seen, the ALSFRS-R trajectories over time are strongly diverging: while one patient remains stable over the course of 12 months, the other patient deteriorates rapidly. This variability matches the variability observed in survival time, and the natural and biological association between the two endpoints [17]: a patient with minimal functional loss has, on average, a longer survival time than a patient with a more aggressive disease trajectory [9]. In the joint modelling framework, we explicitly model this connection between functional loss and the risk of death, and directly account for their interrelationship. As the name suggests, the joint model consists of two models: (1) a model that describes the patient-specific ALSFRS-R trajectory over time, and (2) a model that describes the risk of death based on the modelled ALSFRS-R trajectories.

This process is illustrated in **Fig. 2A**, in which the observed ALSFRS-R scores (*dots*) are used to estimate an underlying longitudinal trajectory for each individual patient (*solid lines*). These modelled trajectories are subsequently used as input, or covariate, in a second model to determine the risk of death (or probability of survival). As a consequence, the patient's ALSFRS-R trajectory alters the instantaneous risk of death over time (**Fig. 2B**), which subsequently changes their probability of survival (**Fig. 2C**). In other words, the joint model links the ALSFRS-R model with the risk function that models survival time. N.B., Cox



models with time-varying covariates are frequently used as alternatives to joint models. Cox models, however, are not appropriate in these settings as they assume that (1) measurement of the longitudinal outcome is not affected by the occurrence of death, (2) the longitudinal outcome remains constant between two visits or until death, and (3) the longitudinal outcome contains minimal measurement error or biological variation [6]. The joint model improves on this by estimating an underlying ‘true’ trajectory for the longitudinal endpoint, estimating its value at the exact time of death and linking this to the risk function for the event of interest.

### **2.3 Defining treatment response and the primary objective**

In RCTs, we can use a joint model to disentangle the effect of treatment [23]. In the VPA study, for example, treatment may reduce the progression rate on ALSFRS-R and, as a consequence, lead to an improvement in survival time. In addition, VPA treatment may also affect important prognostic mechanisms that are not captured by the ALSFRS-R (e.g., weight loss or cognitive decline). As such, treatment may improve survival through two distinct pathways: (1) a pathway that is driven by a change in ALSFRS-R, and (2) a pathway that is driven by a mechanism independent of the ALSFRS-R. From this it follows that we may use the same joint model to address different research questions regarding the effect of treatment, namely:

1. What is the effect of treatment on ALSFRS-R progression?
2. What is the overall effect of treatment on survival time?
3. What is the combined effect of treatment on the ALSFRS-R and survival time, independent of their association?

Importantly, the answer to each of these questions may be different, possibly leading to different conclusions from the same study results. In the following sections, we will provide the joint model output for the VPA study, discuss how these different questions can be answered and how results differ from common conventional methods. The exact research question that is being answered closely coincides with the estimand framework [24], meaning ‘what needs to be estimated to address the research question’. The estimand of an RCT defines, among other things, the primary outcome of interest (e.g., ALSFRS-R, survival, or both) and the strategy for handling events such as death [25]. As we will discuss, the above three questions loosely match the while-alive, treatment policy and composite estimand, respectively [26].

### 3.1 Joint modelling of the VPA study

In **Table 2** we provide an overview of the variables required for the joint model; the data consist of two datasets: (1) ALSFRS-R data and (2) survival information. Defining the joint model starts with a linear mixed model for the ALSFRS-R data. To illustrate, we assume that VPA linearly reduces the ALSFRS-R progression rate, which can be modelled as

$$\eta_i(t) = \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \beta_3 \text{Treatment}_i t$$

$$\beta_{0i} = \beta_0 + \mu_{0i}$$

$$\beta_{1i} = \beta_1 + \mu_{1i}$$

$$\beta_{2i} = 0 + \mu_{2i}$$

Here,  $\eta_i(t)$  is the modelled ALSFRS-R score for the  $i^{\text{th}}$  patient at time  $t$ ,  $\beta_{0i}$  the subject-specific baseline score, and  $\beta_{1i}$  and  $\beta_{2i}$  a subject-specific quadratic trajectory over time (reflecting the individual curves in **Fig. 2A**). In addition,  $\beta_0$  is a common intercept for both

treatment arms (due to randomization); it reflects the average ALSFRS-R score at time of randomization;  $\beta_1$  is the average progression rate in the placebo arm in ALSFRS-R points per month; and  $\beta_3$  is the mean difference in progression rates between treatment arms; it reflects the treatment effect on the ALSFRS-R.

The second step requires a model for the risk of death that contains the ALSFRS-R model as covariate. In addition, we need a term to model any additional effect of treatment on survival that is not captured by the ALSFRS-R. The survival model, or hazard function, can be defined as

$$Hazard_i(Time) = h_0(Time) + \exp[\gamma \cdot Treatment_i + \alpha \cdot \eta_i(t)]$$

where  $h_0(Time)$  is a baseline risk function,  $\gamma$  the additional effect of treatment not captured by the ALSFRS-R, and  $\alpha$  the effect of ALSFRS-R on the risk of death. Similar to a Cox model,  $\gamma$  and  $\alpha$  are simply expressed as log hazard ratio.  $\exp(\alpha)$  indicates how the risk of death changes with each unit increase in ALSFRS-R. It is important to note that  $\eta_i(t)$  corresponds to the ALSFRS-R score predicted by the longitudinal model, and is not the actual observed ALSFRS-R score at time  $t$ . The baseline risk,  $h_0(Time)$ , can be left unspecified similar to a Cox model, but could also be defined using parametric survival distributions. Leaving  $h_0(Time)$  unspecified avoids the need for assumptions, but makes obtaining absolute measures of risk less convenient [6, 15].

After defining the ALSFRS-R and survival sub-models, these are then fitted simultaneously, thereby jointly optimizing all model parameters and adjusting for the interrelationship between ALSFRS-R and survival. As such, model parameters are estimated such that the

model optimally represents both the ALSFRS-R and survival data [6, 27]. This can be achieved by straightforward application of readily available software packages (e.g. the JM or JMBayes package in R [28], the `jmxst`, `merlin` or `stjm` commands in Stata, or the `JMfit` macro in SAS [29]). The **Supplementary Appendix** provides the source code and datasets to replicate the model output for the VPA study presented in **Table 3**. As can be seen, the joint model output confirms the highly significant association between ALSFRS-R and survival time ( $-0.08$  or a HR of  $0.92$ ,  $p < 0.001$ ). In the following sections we will discuss how to address the three research questions outlined above.

### 3.2 Question 1: What is the effect of treatment on ALSFRS-R?

This question coincides with the “while alive” estimand where we are primarily interested in the patient’s functional decline during life. We aim to determine the between-group difference in mean rate of functional loss, as measured by the ALSFRS-R total score, between VPA and placebo, over the follow-up period or until death (whichever occurs first). From the model output in **Table 1** it can be observed that the placebo group declined, on average, with  $0.95$  ALSFRS-R points per month. The time by treatment interaction reflects the mean difference in ALSFRS-R progression rates between placebo and VPA. Although not statistically significant ( $p = 0.34$ ), we can see that the average progression rate in the VPA arm is  $0.11$  points per month faster than placebo ( $-1.06$  vs.  $-0.95$ ; a  $11.6\%$  increase), suggesting a potentially harmful effect of VPA.

The key challenge addressed by the joint model here is that the ALSFRS-R scores for patients who die, do not exist, and are ‘truncated due to death’ [30]. Other methods that ignore the informative missing data often lead to biased results [31]. For example, limiting the analysis to only patients who survived until week 20 ( $N = 24$ , **Fig. 1A**) would result in an

average progression rate of -0.71 ALSFRS-R points per month in the placebo arm and -0.59 in the VPA arm. As the missing data mechanism (death) is associated with the ALSFRS-R, such an analysis not only underestimates the actual progression rate by more than 25%, but also reverses the effect of VPA, which now seems beneficial due to the imbalance in death rate between treatment arms (**Fig. 1B**).

### 3.3 Question 2: What is the overall effect of treatment on survival?

In this setting, we are primarily interested in VPA's overall treatment effect on survival, irrespective of the patient's functional decline, and estimate a "treatment policy" estimand [24]. By substituting  $\eta_i(t)$  in the hazard function (**Section 3.1**), it follows that VPA's overall treatment effect on survival at time  $t$  is the effect mediated through the ALSFRS-R plus the effect not captured by the ALSFRS, or  $\gamma + \alpha\beta_3 t$ , and reflects the log hazard ratio of VPA at a certain time point. Significantly, the hazard ratio of VPA is time-dependent in our model. This is a positive feature of joint models and distinctive from conventional survival models, which assume a constant treatment effect over time [1]. From **Table 1** we can calculate that the hazard ratio of VPA is 1.80 at the beginning of the trial (i.e.,  $\exp(0.59 + -0.08 \cdot [-0.11 \cdot 0])$ ), which increases to 2.15 at month 20 (i.e.,  $\exp(0.59 + -0.08 \cdot [-0.11 \cdot 20])$ ). Thus, the accelerated ALSFRS-R progression rate due to VPA treatment is reflected in the hazard, leading to an increasingly negative effect of VPA on survival over time. Moreover, we can calculate that the effect mediated through the ALSFRS-R is only 23% by month 20 (i.e.,  $-0.08 \cdot -0.11 \cdot 20 / \log(2.15)$ ), suggesting that VPA primarily affects survival through a mechanism not captured by the ALSFRS-R.

Compared to conventional survival models, it seems straightforward that simply comparing the survival curves in **Fig. 1B** should also reflect the overall treatment effect  $T$ , or at least the

average effect on survival during follow-up. If we fit a Cox model, however, we obtain a HR of 1.68 for VPA; 10-20% smaller than estimated by the joint model. This is due to fitting an incorrect model, where the difference between a Cox and joint model depends on the degree of association between the two endpoints [3, 4]. This has important consequences for RCTs: detecting a hazard ratio of 1.68 with 80% power requires 388 patients (assuming a survival probability of 70%), whereas detecting a hazard ratio of 1.80 requires only 302 patients [32]; a difference in sample size of 22.1%.

### 3.4 Question 3: What is the combined effect of treatment on ALSFRS-R and survival?

Finally, in some settings, rather than having particular interest in one endpoint, we could also state two co-primary endpoints without a specific preference for either. In this setting, ‘trial success’ could be defined as a change in either the longitudinal and/or time-to-event endpoint, which coincides with a composite estimand. In a joint model, this could be accommodated by comparing a model with and without the treatment terms. In our example, we could compare a model where we remove the terms  $\beta_3$  and  $\gamma$  (**Table 3**) versus a model that includes these terms. This results in an ANOVA-like p-value, which tests whether the combined effect of treatment improves survival and/or ALSFRS-R, independent of their association (see **Supplementary Appendix** for an example) [9].

This approach is similar to combining both endpoints into a composite endpoint and testing whether this differs between treatment arms. An example is the Combined Assessment of Function and Survival [33]; the patient who died first is ranked worst, whereas the patient who survived and had the least ALSFRS-R loss, ranked best. In contrast to the joint model, however, such endpoints use only one of the two endpoints per patient (if a patient dies, his or her longitudinal data is disregarded). Moreover, if 20% of the patients die at the end of the

trial, 80% of the ranking scores are based on the longitudinal endpoint. This becomes particularly inefficient if treatment affects only one of the two endpoints [9, 26].

#### **4. Final remarks**

In this paper, we have illustrated how joint models can address different research questions in RCTs and have the potential to overcome the limitations of some commonly used conventional methods. Despite the considerable number of publications, and the common co-occurrence of correlated time-to-event and longitudinal data in RCTs, the uptake of joint modelling in actual trials remains negligible [34]. As illustrated in this paper, the interpretation of joint model parameters and those obtained by conventional methods are very similar, can be linked to clinically relevant research questions and, as shown by others, have nowadays been implemented in many software solutions [15].

The main challenges when formulating a joint model are the assumptions required regarding the distributions as well as the associations between the two endpoints [3, 35]. As with any other statistical model, attention will need to be given to potential biases due to the inaccuracy of what we assume, and how this could impact our estimate of the treatment effect [34]. At the design stage of a trial it would be important, for example, to evaluate the impact of changing the trajectory of the longitudinal outcome, the survival pattern, the correlation structure between patients, or how treatment interacts with each endpoint [35].

Many of these considerations start with the exact research question one seeks to answer. Positioning the joint model into the estimand framework, therefore, may better guide one in defining the exact research question, in how best to prespecify the analysis, and in determining which parameter should be of primary interest. This decision process should be a

collaborative effort between experts, statisticians, clinicians, and regulatory scientists to increase mutual understanding of joint models, further mediate their implementation in real-world settings and, ultimately, better utilize the value of joint modelling in clinical trials.

## Acknowledgments

We thank Dr. Kun Jin of the US Food and Drug Administration for discussions and for providing insight into the regulatory considerations for joint models, and Brenda Vollers-King for assistance with language editing.

## REFERENCES

- [1] van Oudenhoven FM, Swinkels SHN, Hartmann T, Soininen H, van Hees AMJ, Rizopoulos D. Using joint models to disentangle intervention effect types and baseline confounding: an application within an intervention study in prodromal Alzheimer's disease with Fortasyn Connect. *BMC Med Res Methodol*. 2019;19:163.
- [2] Ediebah DE, Galindo-Garre F, Uitdehaag BM, Ringash J, Reijneveld JC, Dirven L, et al. Joint modeling of longitudinal health-related quality of life data and survival. *Qual Life Res*. 2015;24:795-804.
- [3] Chen LM, Ibrahim JG, Chu H. Sample size and power determination in joint modeling of longitudinal and survival data. *Stat Med*. 2011;30:2295-309.
- [4] Ibrahim JG, Chu H, Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *J Clin Oncol*. 2010;28:2796-801.
- [5] Lawrence Gould A, Boye ME, Crowther MJ, Ibrahim JG, Quartey G, Micallef S, et al. Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group. *Stat Med*. 2015;34:2181-95.



- [6] Rizopoulos D. Joint models for longitudinal and time-to-event data : with applications in R. Boca Raton: CRC Press; 2012.
- [7] Levine SZ, Goldberg Y, Samara M, Davis JM, Leucht S. Joint modeling of dropout and outcome in three pivotal clinical trials of schizophrenia. *Schizophr Res*. 2015;164:122-6.
- [8] Garcia-Hernandez A, Perez T, Pardo MDC, 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*. 2020;19:909-27.
- [9] van Eijk RPA, Eijkemans MJC, Rizopoulos D, van den Berg LH, Nikolakopoulos S. Comparing methods to combine functional loss and mortality in clinical trials for amyotrophic lateral sclerosis. *Clin Epidemiol*. 2018;10:333-41.
- [10] Deslandes E, Chevret S. Assessing surrogacy from the joint modelling of multivariate longitudinal data and survival: application to clinical trial data on chronic lymphocytic leukaemia. *Stat Med*. 2007;26:5411-21.
- [11] Spertus JV, Hatfield LA, Cohen DJ, Arnold SV, Ho M, Jones PG, et al. Integrating Quality of Life and Survival Outcomes in Cardiovascular Clinical Trials. *Circ Cardiovasc Qual Outcomes*. 2019;12:e005420.
- [12] Sudell M, Kolamunnage-Dona R, Tudur-Smith C. Joint models for longitudinal and time-to-event data: a review of reporting quality with a view to meta-analysis. *BMC Med Res Methodol*. 2016;16:168.
- [13] Asgari S, Khalili D, Zayeri F, Azizi F, Hadaegh F. Dynamic prediction models improved the risk classification of type 2 diabetes compared with classical static models. *J Clin Epidemiol*. 2021;140:33-43.
- [14] Farcomeni A, Pareek B, Ghosh P. Discussion on 'Joint modeling of survival and longitudinal non-survival data' by Gould et al. *Stat Med*. 2015;34:2198-9.

- [15] Gould AL, Boye ME, Crowther MJ, Ibrahim JG, Quartey G, Micallef S, et al. Responses to discussants of 'Joint modeling of survival and longitudinal non-survival data: current methods and issues. report of the DIA Bayesian joint modeling working group'. *Stat Med*. 2015;34:2202-3.
- [16] Piepers S, Veldink JH, de Jong SW, van der Tweel I, van der Pol WL, Uijtendaal EV, et al. Randomized sequential trial of valproic acid in amyotrophic lateral sclerosis. *Ann Neurol*. 2009;66:227-34.
- [17] Westeneng HJ, Debray TPA, Visser AE, van Eijk RPA, Rooney JPK, Calvo A, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. *Lancet Neurol*. 2018;17:423-33.
- [18] van Eijk RPA, Nikolakopoulos S, Roes KCB, Kendall L, Han SS, Lavrov A, et al. Challenging the Established Order: Innovating Clinical Trials for Amyotrophic Lateral Sclerosis. *Neurology*. 2021.
- [19] Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci*. 1999;169:13-21.
- [20] European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of amyotrophic lateral sclerosis. [https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-amyotrophic-lateral-sclerosis\\_en.pdf](https://www.ema.europa.eu/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-amyotrophic-lateral-sclerosis_en.pdf); 2016. p. Accessed 19 Nov 2021.
- [21] Food Drug Administration Center for Drugs Evaluation Research. Guidance for Industry: Amyotrophic Lateral Sclerosis: Developing Drugs for Treatment <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM596718.pdf>; 2021. p. Accessed 4 May 2020.

- [22] van Eijk RPA, Nikolakopoulos S, Ferguson TA, Liu D, Eijkemans MJC, van den Berg LH. Increasing the efficiency of clinical trials in neurodegenerative disorders using group sequential trial designs. *J Clin Epidemiol*. 2018;98:80-8.
- [23] van Oudenhoven FM, Swinkels SHN, Ibrahim JG, Rizopoulos D. A marginal estimate for the overall treatment effect on a survival outcome within the joint modeling framework. *Stat Med*. 2020;39:4120-32.
- [24] ICH E9 (R1). Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials.  
[https://database.ich.org/sites/default/files/E9-R1\\_Step4\\_Guideline\\_2019\\_1203.pdf](https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf); 2019. p. Accessed 20 Nov 2021.
- [25] Oude Rengerink K, Mitroiu M, Teerenstra S, Petavy F, Roes KCB. Rethinking the intention-to-treat principle: one size does not fit all. *J Clin Epidemiol*. 2020;125:198-200.
- [26] van Eijk RPA, Roes KCB, de Greef-van der Sandt I, van den Berg LH, Lu Y. Functional loss and mortality in randomized clinical trials for amyotrophic lateral sclerosis: to combine, or not to combine - that is the estimand. *Clin Pharmacol Ther*. 2022.
- [27] Wulfsohn MS, Tsiatis AA. A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics*. 1997;53:330-9.
- [28] Alsefiri M, Sudell M, Garcia-Finana M, Kolamunnage-Dona R. Bayesian joint modelling of longitudinal and time to event data: a methodological review. *BMC Med Res Methodol*. 2020;20:94.
- [29] Chesnaye NC, Tripepi G, Dekker FW, Zoccali C, Zwinderman AH, Jager KJ. An introduction to joint models-applications in nephrology. *Clin Kidney J*. 2020;13:143-9.
- [30] Colantuoni E, Scharfstein DO, Wang C, Hashem MD, Leroux A, Needham DM, et al. Statistical methods to compare functional outcomes in randomized controlled trials with high mortality. *BMJ*. 2018;360:j5748.

- [31] Olivier T, Haslam A, Prasad V. Informative censoring due to missing data in quality of life was inadequately assessed in most oncology randomized controlled trials. *J Clin Epidemiol*. 2021;139:80-6.
- [32] Schoenfeld DA, Richter JR. Nomograms for calculating the number of patients needed for a clinical trial with survival as an endpoint. *Biometrics*. 1982;38:163-70.
- [33] Berry JD, Miller R, Moore DH, Cudkowicz ME, van den Berg LH, Kerr DA, et al. The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials. *Amyotroph Lateral Scler Frontotemporal Degener*. 2013;14:162-8.
- [34] Price DL, Wang Y. Commentary on 'Joint modeling of survival and longitudinal non-survival data: current methods and issues'. *Stat Med*. 2015;34:2200-1.
- [35] Xu J, Psioda MA, Ibrahim JG. Bayesian design of clinical trials using joint models for longitudinal and time-to-event data. *Biostatistics*. 2020.

**Table 1.** Uptake of joint modelling in 61 consecutive clinical trials.

<b>Characteristic</b>	<b>No. of clinical trials</b> <i>(N = 61)</i>
<b>Primary endpoint of clinical trial</b>	
Time-to-event	49 (80%)
Longitudinal	12 (20%)
<b>Time-to-event analysis</b>	
Logrank test	26 (43%)
Stratified logrank test	16 (26%)
Cox model with covariate adjustment	17 (28%)
Mediation analysis with longitudinal outcome	2 (3%)
<b>Longitudinal analysis</b>	
Mixed model for repeated measures	25 (41%)
Linear mixed model	11 (18%)
Simple regression	22 (36%)
Composite endpoint with time-to-event	1 (2%)
Not reported	2 (3%)
<b>Joint modelling of time-to-event and longitudinal data,</b> <i>Yes</i>	0 (0%)

Overview of statistical strategies applied in consecutively randomized clinical trials, published between January 1, 2019 and December 31, 2019 in the NEJM, Lancet and JAMA, that reported the results on both endpoints in the publication; study selection is described in the **Supplementary Appendix**.

**Table 2.** Summary of the datasets required to fit a joint model in the VPA study

<b>Dataset</b>	<b>Description</b>
1. Longitudinal dataset	Number of rows is equal to sample size times number of visits
i. Subject ID	Identifier for individual patient
ii. Visiting time	Time of visit in months from baseline, options: 0, 2, 4, 8, 12, 16, 20
iii. ALSFRS-R total score	ALSFRS-R total score at visit
iv. Treatment	Treatment indicator, options: placebo, valproic acid
2. Time-to-event dataset	Number of rows is equal to sample size
i. Subject ID	Identifier for individual patient
ii. Survival time	Time in months between baseline and death or last follow-up
iii. Vital status	Status indicator, options: dead, alive
iv. Treatment	Treatment indicator, options: placebo, valproic acid

**Table 3.** Joint model output of the Valproic Acid study.

Parameter	Estimate	95% CI	P-value
<i>ALSFRS-R model</i>			
Intercept ( $\beta_0$ )	40.2	39.3 to 41.1	<0.001
Time ( $\beta_1$ )	-0.95	-1.11 to -0.79	<0.001
Time x Treatment ( $\beta_3$ )	-0.11	-0.33 to 0.11	0.34
<i>Survival model</i>			
Log HR Treatment ( $\gamma$ )	0.59	-0.12 to 1.29	0.10
Log HR ALSFRS-R ( $\alpha$ )	-0.08	-0.12 to -0.04	<0.001

The exact model specifications are given in the **Supplementary Appendix**. Abbreviations:

ALSFRS-R = Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; HR = hazard ratio.

**Figure 1 title.** Functional loss and survival in a randomized controlled clinical trial for ALS.

**Figure 1 legend.** The observed individual ALSFRS-R trajectories and the Kaplan-Meier curves of each treatment arm in a randomized, placebo-controlled trial with valproic acid in patients with ALS [16]. The original trial randomized 163 patients; as the trial stopped early, nine patients, with less than 4 weeks of follow-up time, were censored administratively and excluded from the current analysis.

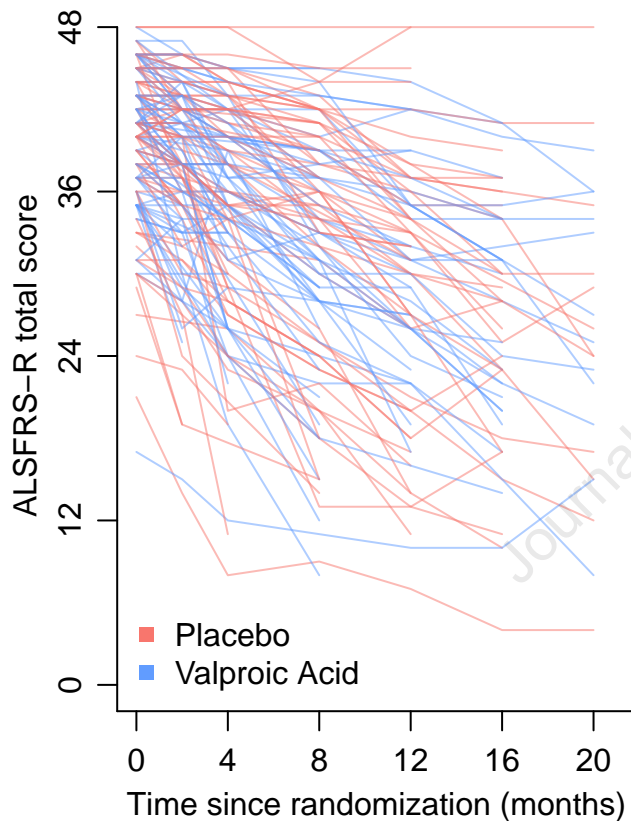
**Figure 2 title.** The joint modelling framework illustrated for amyotrophic lateral sclerosis.

**Figure 2 legend.** (A) Joint models estimate an underlying trajectory (*solid line*) based on the observed longitudinal outcomes (*dots*). These modelled trajectories are then linked with the patient's hazard for the event (B), and thereby directly associated with his or her probability of survival during follow-up (C).



**A. Longitudinal outcome**

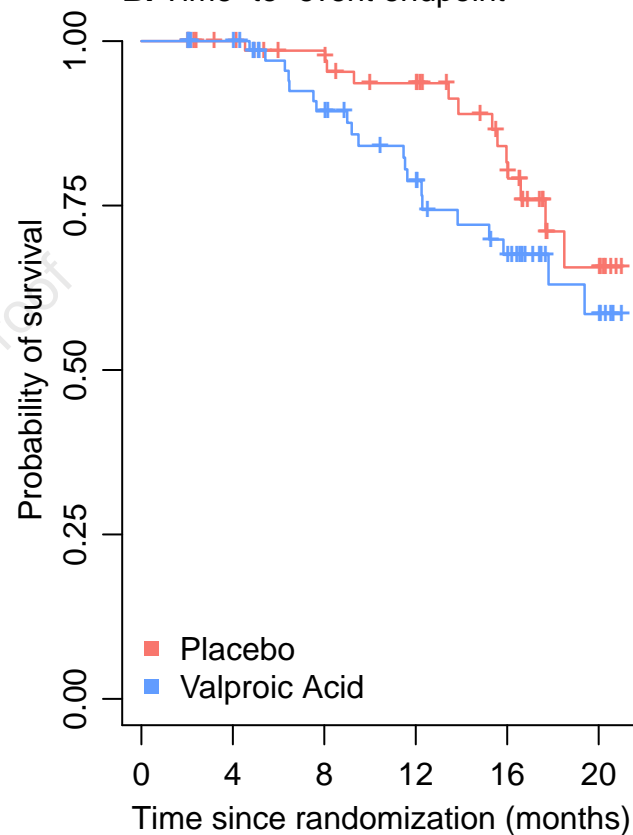
Journal Pre-proof



**No. patients**

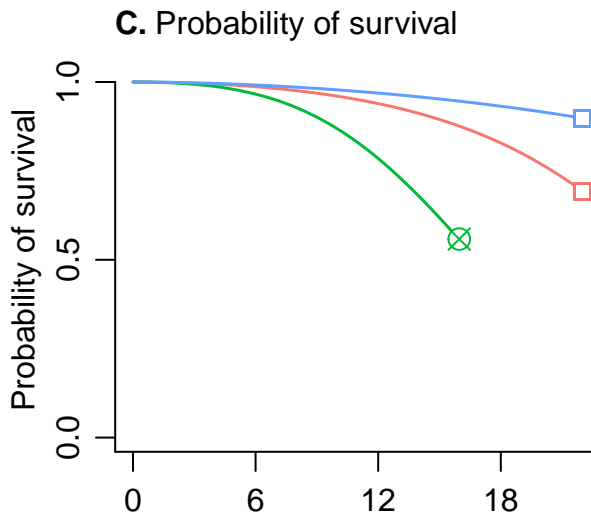
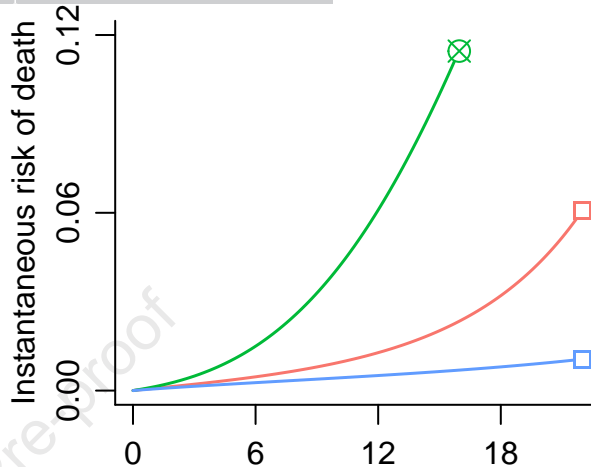
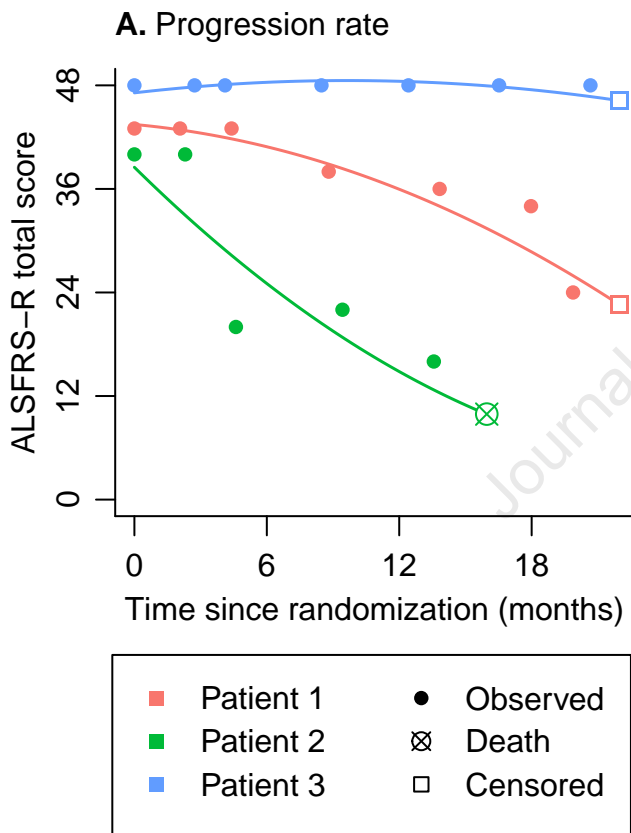
Placebo	77	64	60	50	29	12
VPA	77	65	52	36	26	12

**B. Time-to-event endpoint**



**No. at risk**

Placebo	77	70	65	53	33	12
VPA	77	72	58	44	29	13



**HIGHLIGHTS**

- Time-to-event and longitudinal endpoints commonly co-occur in clinical trials
- Joint modelling is a powerful approach to better understand the treatment effect
- Joint models may help to address informative censoring in longitudinal endpoints
- Efficiency gains can be achieved by simultaneous assessment of all available data
- Limitations are increased complexity and the risk of defining an erroneous model