Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
0(0) 1–12

DOI: 10.1177/0962280218802300
journals.sagepub.com/home/smm
SAGE

Joint modeling of multiple repeated measures and survival data using multidimensional latent trait linear mixed model

Jue Wang[1] and Sheng Luo[2]

Abstract

Impairment caused by Amyotrophic lateral sclerosis (ALS) is multidimensional (e.g. bulbar, fine motor, gross motor) and progressive. Its multidimensional nature precludes a single outcome to measure disease progression. Clinical trials of ALS use multiple longitudinal outcomes to assess the treatment effects on overall improvement. A terminal event such as death or dropout can stop the follow-up process. Moreover, the time to the terminal event may be dependent on the multivariate longitudinal measurements. In this article, we develop a joint model consisting of a multidimensional latent trait linear mixed model (MLTLMM) for the multiple longitudinal outcomes, and a proportional hazards model with piecewise constant baseline hazard for the event time data. Shared random effects are used to link together two models. The model inference is conducted using a Bayesian framework via Markov chain Monte Carlo simulation implemented in Stan language. Our proposed model is evaluated by simulation studies and is applied to the Ceftriaxone study, a motivating clinical trial assessing the effect of ceftriaxone on ALS patients.

Keywords

Amyotrophic lateral sclerosis, informative dropout, longitudinal data, Markov chain Monte Carlo, mixed model

1 Introduction

Amyotrophic lateral sclerosis (ALS), sometimes referred to as Lou Gehrig's Disease, is a neurodegenerative disorder characterised by progressive muscular paralysis reflecting degeneration of motor neurones in the primary motor cortex, corticospinal tracts, brainstem and spinal cord.[1] ALS is the most common motor neuron disease and the majority of ALS patients die of respiratory failure within two to five years of clinical onset. Typically, 30,000 Americans have the disease at any given time. People of all races and ethnic backgrounds are affected. The pathogenesis of ALS remains largely unknown and there is no cure, while many clinical trials are conducted to search for neuroprotective treatments capable of extending the length and meaningful quality of life for patients (e.g. the study of Ceftriaxone,[2] the PRO-ACT database,[3] and the study of MCI-186[4]). ALS causes impairment in multiple domains (e.g. bulbar, motor, and respiratory). The disease progresses heterogeneously in time and across domains: decline may be observed in some, but not all health outcomes at any given time interval and the trajectory of progression may vary between different domains. Therefore, no single health outcome reliably reflects the full spectrum of ALS severity and progression.

Commonly used endpoints in ALS trials include assessments of muscle strength, respiratory function, and mortality.[5] However, these assessments require specialized equipment to administer and can be time-consuming. In contrast, the ALS Functional Rating Scale (ALSFRS) provides a physician-generated estimate of the patient's degree of functional impairment, evaluating multiple clinical aspects of ALS, including bulbar,

[1]Department of Biostatistics, The University of Texas Health Science Center at Houston, Houston, TX, USA
[2]Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, USA

Corresponding author:
Sheng Luo, Department of Biostatistics and Bioinformatics, Duke University Medical Center, 2400 Pratt St, 7040 North Pavilion, Durham 27705, NC, USA.
Email: sheng.luo@duke.edu

fine motor, gross motor function, and respiratory disability.[6] The ALSFRS includes 10 questions (Web Table 1) rating patients' level of functional impairment in performing common tasks, e.g. Item 1 (speech) and Item 5 (cutting food). Each task is rated on a five-point scale from 0 (unable to attempt the task) to 4 (normal ability) at multiple visits during the ALS trials, leading to the multivariate longitudinal data structure. The summated ALSFRS total score (sum of the scores of all 10 ordinal ALSFRS items, with the range of 0–40) has been frequently used to evaluate patients' overall disease severity.[7–9] However, researchers are increasingly suggesting that the total score (derived by summing multiple ordinal items) is a relatively imprecise indicator of underlying disease severity because it ignores the differences among various response patterns that result in identical total score.[10,11] To this end, Bacci et al.[5] applied Rasch[12] and Graded Response Model[13] under the item response theory (IRT) framework to each item of the ALSFRS, which provides information on the scaling metrics and item-level performance of an instrument, as well as specific characteristics of the scale. To account for the complex correlation structure in multivariate longitudinal data, an extension of the IRT framework, the multilevel IRT model has been developed.[14–16] A common feature of the multilevel IRT model is that a univariate latent variable representing the unobserved true disease severity is regressed on predictors (e.g. treatment and time) and subject-specific random effects (describing the between-subject differences), where the multiple items are viewed as clinical manifestations of the latent variable. This unidimensional assumption has been relaxed in Wang and Luo[17] by developing a multidimensional latent trait linear mixed model (MLTLMM) to allow multiple latent variables and within-item multidimensionality (one outcome can be a manifestation of more than one latent variable) in the analysis of the multivariate longitudinal data.

During the follow-up process of ALS studies, some patients may have terminal events such as death, dropout, and withdrawal of consent. The terminal events and the multivariate longitudinal measurements are often not independent. For example, more severe ALS (as manifested by the ALSFRS) often increases the risk of death, which in turn makes any subsequent ALSFRS measurement impossible. The presence of such a dependent failure time is often termed as ''informative dropout,'' ignoring which could lead to biased estimates.[18] To address this issue, joint analysis of survival with repeated measures has been increasingly common.[19–22] In this article, we develop a joint modeling framework in which a MLTLMM that allows multiple latent variables and within-item multidimensionality is used for multivariate longitudinal outcomes and a proportional hazards model with piecewise constant baseline hazard is used for the dependent terminal event. The two models are linked via shared random effects. The rest of the article proceeds as follows. In Section 2, we describe a motivating clinical trial. Section 3 discusses the joint MLTLMM model, Bayesian inference, and Bayesian model selection criteria. Section 4 provides a simulation study to assess the performance of the proposed model. In Section 5, we apply the proposed model to the motivating clinical trial dataset. Section 6 gives some concluding remarks and discussions.

2 A motivating clinical trial

The methodological development is motivated by the Ceftriaxone study,[2] a double-blind, placebo controlled multi-phase clinical trial (completed in July 2012) with 513 subjects to determine if ceftriaxone slows the disease progression of ALS. The revised ALS Functional Rating Scale (ALSFRS-R) is the primary outcome in the Ceftriaxone study, which includes the 10 original ALSFRS items used in Wang and Luo[17] and two additional items. To obtain comparable results, the two additional items are excluded. In the Ceftriaxone study, ALSFRS was measured at 21 visits (screening, week 4, week 16, and every eight weeks starting from week 16 to week 152, plus a final study visit performed when subject completes trial), leading to multivariate longitudinal data structure (multiple ALSFRS items were repeatedly measured). Five subjects are excluded because of missing ALSFRS or the variable of forced vital capacity (FVC) at baseline and the analysis is based on the remaining 508 patients (172 in placebo and 336 in Ceftriaxone). In Wang and Luo,[17] we found three ALS impaired domains among the 10 items, i.e. bulbar function (including items of speech, salivation, and swallowing), fine motor function (including items of handwriting, cutting, and dressing), and gross motor function (including items of walking and climbing), while Item 7 (turning) overlaps on the fine motor and gross motor functions and Item 10 (breathing) slightly loads on the bulbar function.

Before the end of the study, some patients (86 and 164 patients in the placebo and Ceftriaxone groups, respectively) died. Time to death may be correlated with the ALSFRS because the more severe ALS, patients (with lower ALSFRS scores) have increased risk of death. To visualize the correlation between the terminal event and the outcome measurement, Figure 1 shows Item 2 (salivation) score over time for each patient who died with over 1.5 years of follow-up (first panel), who was censored with over 1.5 years of follow-up (second panel), who

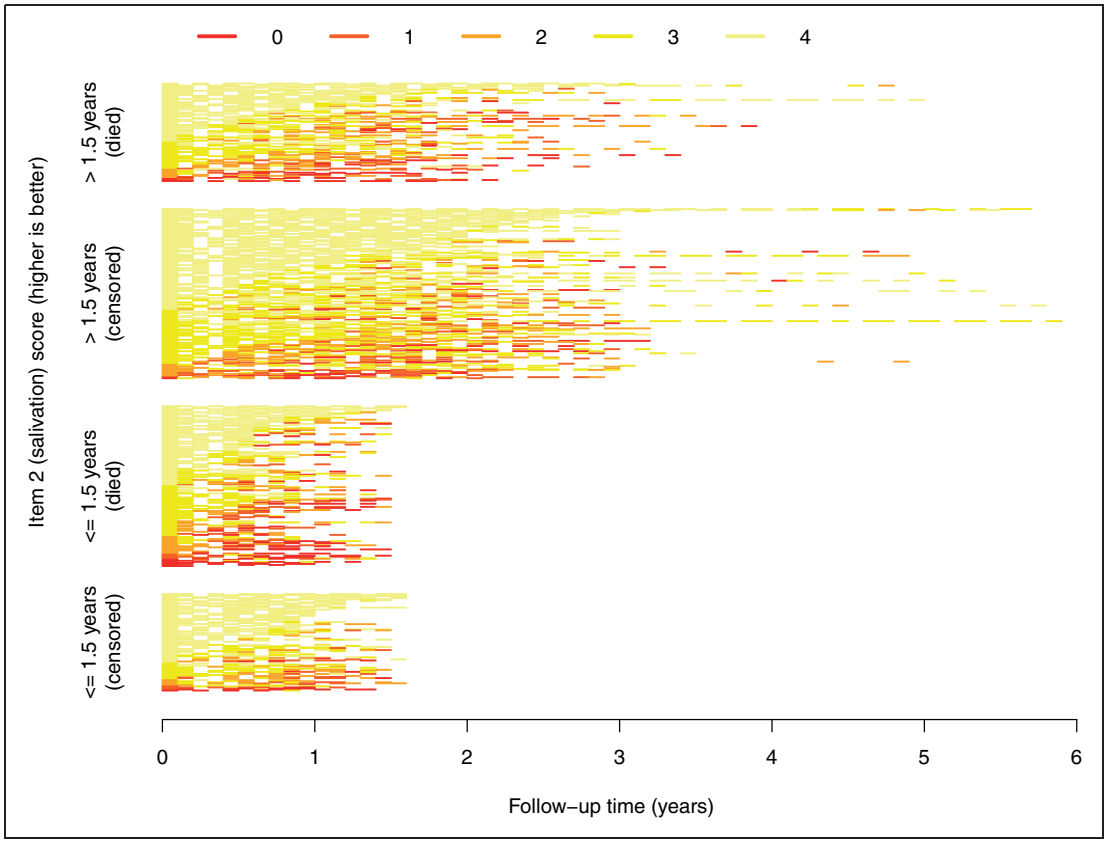

Figure 1. Item 2 (salivation) score over time for each patient.

died with less than or equal to 1.5 years of follow-up (third panel), and who was censored with less than or equal to 1.5 years of follow-up (fourth panel). At any time point, patients with longer survival time tend to have higher salivation score (less severe ALS) compared with those with shorter survival time (first panel vs. third panel), suggesting that patients with longer survival times may have better health condition and slower disease progression. Similar patterns are observed in other items. To this end, we will develop a joint modeling framework with multidimensional latent trait model to account for the disease impairment in multiple domains and proportional hazards model for time to death.

3 Model and estimation

3.1 Model Formulation and Likelihood

In the context of clinical trials with multiple outcomes, the data structure is often of the type $\{y_{ik}(t_{ij}), t_i, \delta_i\}$, where $y_{ik}(t_{ij})$ is the kth $(k = 1, \ldots, K)$ outcome, which can be binary, ordinal, or continuous, for subject i $(i = 1, \ldots, I)$ at visit j $(j = 1, \ldots, J_i)$ recorded at time t_{ij} from the study onset, $t_i = min(T_i^*, C_i)$ is the observed event time to functional disability, as the minimum between the true event time T_i^* and the censoring time C_i which are assumed to be independent, and δ_i is the censoring indicator (1 if the event is observed, and 0 otherwise). To illustrate the MLTLMM modeling framework, we assume that there are P (with $P < K$) latent variables (LVs) representing the underlying disease severity scores and denote them as $\boldsymbol{\theta}_i(t) = (\theta_i^{(1)}(t), \ldots, \theta_i^{(p)}(t), \ldots, \theta_i^{(P)}(t))'$ for subject i at time t, where the superscript (p) $(p = 1, \ldots, P)$ denotes the pth latent variable. We introduce the first level MLTLMM model for continuous outcomes, binary outcomes, and the cumulative probabilities of ordinal outcomes

$$\begin{aligned} y_{ik}(t) &= a_k + \boldsymbol{b}_k'\boldsymbol{\theta}_i(t) + \varepsilon_{ik}(t) \\ \text{logit}\{p(y_{ik}(t) = 1|\boldsymbol{\theta}_i(t))\} &= a_k + \boldsymbol{b}_k'\boldsymbol{\theta}_i(t) \\ \text{logit}\{p(y_{ik}(t) \le l\,|\boldsymbol{\theta}_i(t))\} &= a_{kl} - \boldsymbol{b}_k'\boldsymbol{\theta}_i(t) \end{aligned} \tag{1}$$

where a_k and $\boldsymbol{b}_k = (b_k^{(1)}, \ldots, b_k^{(P)})'$ are the parameters specific to outcome k, the random errors $\varepsilon_{ik}(t) \sim N(0, \sigma_{\varepsilon_k}^2)$ are independent and identically distributed, and $l = 1, 2, \ldots, n_k - 1$ is the lth level of the kth outcome, which is ordinal with n_k categories. A major feature of these models is that they all incorporate $\boldsymbol{\theta}_i(t)$ and explicitly combine information from all outcomes. This is one of the simplest ways to conceptualize the disease severity scores that allow the overall treatment effects to be defined. To model the dependence of severity scores $\boldsymbol{\theta}_i(t)$ on covariates, we propose the second level multivariate linear mixed model (LMM)

$$\theta_i^{(p)}(t) = \boldsymbol{\beta}^{(p)'} \boldsymbol{X}_i^{(p)}(t) + \boldsymbol{u}_i^{(p)'} \boldsymbol{Z}_i^{(p)}(t) + e_i^{(p)}(t) \tag{2}$$

where $\boldsymbol{X}_i^{(p)}(t)$ and $\boldsymbol{Z}_i^{(p)}(t)$ are the covariates corresponding to fixed and random effects, respectively, for each latent variable $\theta_i^{(p)}(t)$. They can include covariates of interest such as treatment and time. The vector $\boldsymbol{u}_i = (\boldsymbol{u}_i^{(1)}, \ldots, \boldsymbol{u}_i^{(P)})'$ contains all the random effects for subject i, which are assumed to be normally distributed as $N_P(0, \Sigma)$, where Σ is the covariance matrix. The correlation among the latent variables are accounted for by the correlation among the elements in $\boldsymbol{u}_i$. The residuals $e_i^{(p)}(t)$ are assumed to be independent from $\boldsymbol{u}_i$ and $e_i^{(p)}(t) \sim N(0, \sigma_e^{(p)^2})$.

For notational convenience, we let $\boldsymbol{a} = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_k', \ldots, \boldsymbol{a}_K')'$, with $\boldsymbol{a}_k$ being numeric for binary and continuous outcomes and $\boldsymbol{a}_k = (a_{k1}, \ldots, a_{kn_k-1})'$ for the kth ordinal outcome with n_k categories. We let $\boldsymbol{b} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K)'$, which is the K by P latent factor loading matrix. Because the MLTLMM model is over-parameterized, additional constraints are required to make it identifiable, which has been discussed in detail in Wang and Luo.[17]

Let the parameter vector be $\Theta = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\beta}, \Sigma, \sigma_{\epsilon_k}, \sigma_e^{(p)}\}$. Conditional on the random effects $\boldsymbol{u}_i$, all the measurements of each subject are assumed to be independent. We have the full likelihood of multiple longitudinal outcomes for subject i as follows

$$L_y(\Theta, \boldsymbol{u}_i; \boldsymbol{y}_i) = \left[\prod_{j=1}^{J_i} \prod_{k=1}^{K} p(y_{ik}(t_{ij}) | \boldsymbol{u}_i) \right] p(\boldsymbol{u}_i | \Sigma) \tag{3}$$

where $p(\boldsymbol{u}_i|\Sigma)$ is the density function of random effects vector $\boldsymbol{u}_i$.

To model the survival process, we use the proportional hazards model

$$h_i(t) = h_0(t) \exp\{\boldsymbol{W}_i \boldsymbol{\gamma} + \boldsymbol{v}' \boldsymbol{u}_i\} \tag{4}$$

where $\boldsymbol{v}$ measure the association between the two models linked together via the shared random effects $\boldsymbol{u}_i$. Although some covariates in the covariate vector $\boldsymbol{W}_i$ may overlap with $\boldsymbol{X}_i(t)$, $\boldsymbol{W}_i$ does not include any time-varying covariates. We adopt a piecewise constant function to approximate the baseline hazard function $h_0(t)$. It has been shown that survival models using a piecewise constant baseline hazard function yield good estimators for both fixed effects and frailty,[23,24] although fixed cut points need to be specified a priori. It is more flexible than assuming a particular parametric baseline hazard distribution (e.g. a Weibull distribution) and it retains enough model structure.[25] Specifically, given a set of fixed time points $\tau_1, \tau_2, \ldots, \tau_Q$ and $\tau_0 = 0$ or the smallest failure time and the baseline hazard vector $\boldsymbol{g} = (g_1, g_2, \ldots, g_Q)'$, we define the piecewise constant baseline hazard function as $h_0(t) = \sum_{q=1}^{Q} g_q I_q(t)$, with indicator function $I_q(t) = 1$ if $\tau_{q-1} \leq t < \tau_q$ and 0 otherwise. The likelihood of observing event outcome t_i and δ_i for subject i is

$$L_s(\Theta_s; t_i, \delta_i, \boldsymbol{u}_i) = h_i(t_i)^{\delta_i} S_i(t_i) \tag{5}$$

where the survival function $S_i(t_i) = \exp\{-\int_0^{t_i} h_i(s)\mathrm{d}s\}$ and the parameter vector for the survival process is $\Theta_s = (\boldsymbol{\gamma}', \boldsymbol{v}')'$. Note that the shared random-effects formulation in the survival model (4) leads to a closed-form expression for the survival function $S_i(t_i)$ which provides a great computational advantage. In addition, we consider a special case of the model where the occurrence of the terminal event is independent to the longitudinal outcomes (i.e. $\boldsymbol{v} = \boldsymbol{0}$) and we refer to it as the reduced model (RM).

3.2 Bayesian inference and model selection

To make inference on the parameter vector Θ, we use Bayesian methods based on Markov chain Monte Carlo (MCMC) posterior simulations. The prior distributions of parameters a_k of the continuous outcomes is $a_k \sim N(0, 100)$, with mean 0 and variance 100. To obtain the prior distributions for the threshold parameters of ordinal outcome k, we let $a_{k1} \sim N(0, 100)$, and $a_{kl} = a_{k,l-1} + \Delta_l$ for $l = 2, \ldots, n_k - 1$, with $\Delta_l \sim N(0, 100)I(0,)$, i.e.

normal distribution left truncated at 0. Prior distributions for all elements in $\boldsymbol{b}$ (except for constrained parameters, see Wang and Luo[17]), $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ are $N(0, 100)$. We use the prior distribution Uniform $[-1, 1]$ for all the correlation coefficients ρ in the covariance matrix Σ, and Inverse-Gamma $(0.01, 0.01)$ for all variance and covariance parameters. The model fitting is implemented in Stan (version 2.17.0),[26] which is a probabilistic programming language implementing statistical inference. Note that Stan requires variable types to be declared prior to modeling. The declaration of matrix Σ as a covariance matrix ensures it to be positive-definite by rejecting the samples that cannot produce positive-definite matrix Σ. Multiple chains with over-dispersed initial values are run to analyze data and the Gelman–Rubin diagnostic[27] is used to ensure that the scale reduction $\hat{R}$ of all parameters are smaller than 1.1. Moreover, we use the trace plots and autocorrelation functions to ensure the chain convergence. To facilitate easy reading and implementation of the proposed model, the Stan code is given in the Web Supplemental Material.

There are a wide variety of model selection criteria in Bayesian inference. Because of the mixture framework in our model, we use the DIC_3 measurement.[28] The DIC_3 is defined as $\mathrm{DIC}_3 = \overline{D(\Theta)} + \tau_D$, where $\overline{D(\Theta)} = -2\mathrm{E}_{\Theta|\boldsymbol{D}}\left\{\log\left[\prod_{i=1}^{I} f(\boldsymbol{y}_i, t_i|\Theta)\right]\right\}$ is the posterior mean deviance, $\tau_D = \overline{D(\Theta)} + 2\log\left\{\mathrm{E}_{\Theta|\boldsymbol{D}}\left[\prod_{i=1}^{I} f(\boldsymbol{y}_i, t_i|\Theta)\right]\right\}$ is a measure of the effective number of parameters in the model, and $\mathrm{E}_{\Theta|\boldsymbol{D}}(.)$ is the expectation with respect to the joint posterior distribution $\pi(\Theta|\boldsymbol{D})$, $\boldsymbol{D}$ denoting the observed data. Thus, we have $\mathrm{DIC}_3 = -4\mathrm{E}_{\Theta|\boldsymbol{D}}\left\{\log\left[\prod_{i=1}^{I} f(\boldsymbol{y}_i, t_i|\Theta)\right]\right\} + 2\log\left\{\mathrm{E}_{\Theta|\boldsymbol{D}}\left[\prod_{i=1}^{I} f(\boldsymbol{y}_i, t_i|\Theta)\right]\right\}$. Applying Monte Carlo approximation

$$\widehat{\mathrm{DIC}}_3 = -\frac{4}{M}\sum_{m=1}^{M}\sum_{i=1}^{I}\log\left\{f(\boldsymbol{y}_i, t_i|\Theta^{(m)})\right\} + 2\log\left\{\frac{1}{M}\sum_{m=1}^{M}\prod_{i=1}^{I} f(\boldsymbol{y}_i, t_i|\Theta^{(m)})\right\}$$

where $\Theta^{(m)}$ is the mth ($m = 1, \ldots, M$) post burn-in posterior samples of parameter vector Θ. A smaller value of DIC_3 indicates a better-fitting model. In addition, we use the expected Akaike information criterion (EAIC) and the expected Bayesian (or Schwarz) information criterion (EBIC) as model selection tools.[29] The EAIC and EBIC can be estimated as $\mathrm{EAIC} = \overline{D(\Theta)} + 2p$ and $\mathrm{EBIC} = \overline{D(\Theta)} + p\log(I)$, where p is the number of parameters in the parameter vector Θ, and I is the number of individuals. Smaller values of EAIC and EBIC indicate better predictive ability of the model.

4 Simulation studies

In this section, we conduct an extensive simulation study with two settings to investigate the performance of the proposed joint model (JM) and the simpler reduced model (RM). Specifically, we generate 200 datasets with sample size $N = 600$ subjects and seven visits (baseline and six follow-up visits, $J_i = 7$) for each subject. The data structure is similar to the motivating Ceftriaxone dataset, and it has $k = 1, \ldots, 10$ ordinal outcomes with five categories each.

Specifically, we generate data by the following model

$$\theta_i^{(p)}(t_{ij}) = \beta_0^{(p)} + \beta_1^{(p)}x_i + \beta_2^{(p)}t_{ij} + \beta_3^{(p)}x_i t_{ij} + u_{i0}^{(p)} + u_{i1}^{(p)}t_{ij} + e_i^{(p)}(t_{ij}),$$
$$h_i(t) = h_0(t)\exp\left[\gamma x_i + \sum_{p=1}^{P}\left(\nu_1^{(p)}u_{i0}^{(p)} + \nu_2^{(p)}u_{i1}^{(p)}\right)\right]$$

where $P = 2$, and the covariate x_i takes value 0 or 1 each with probability 1/2 to mimic the treatment assignment. The time vector $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{i7})' = (0, 1, 2, 3, 4, 5, 6)'$. We set the regression coefficients to be $\boldsymbol{\beta}^{(1)} = \left(\beta_0^{(1)}, \beta_1^{(1)}, \beta_2^{(1)}, \beta_3^{(1)}\right)' = (0.5, -1, 0.5, -0.4)'$ and $\boldsymbol{\beta}^{(2)} = \left(\beta_0^{(2)}, \beta_1^{(2)}, \beta_2^{(2)}, \beta_3^{(2)}\right)' = (1, -0.5, 1, -1)'$. For simplicity, we use a constant baseline hazard function $h_0(t) = h_0 = 0.1$ and coefficient $\gamma = -1$. The random effects vector $\boldsymbol{u}_i = \left(u_{i0}^{(1)}, u_{i1}^{(1)}, u_{i0}^{(2)}, u_{i1}^{(2)}\right)'$ is simulated from a multivariate normal distribution with mean 0 and covariance matrix Σ generated with the following parameters: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 1$, $\rho_{12} = 0.4$, $\rho_{13} = 0.2$, $\rho_{14} = 0.2$, $\rho_{23} = 0.1$, $\rho_{24} = 0.2$, and $\rho_{34} = 0.5$. Random errors $e_i^{(p)}(t_{ij}) \sim N(0, \sigma_e^{2(p)})$ with $\sigma_e^{(1)} = \sigma_e^{(2)} = 0.5$. For identifiability, we set $a_{11} = a_{21} = 0$, $b_1^{(1)} = b_2^{(2)} = 1$, and $b_1^{(2)} = 0$, while the values of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are listed in Web Tables 2 and 3.

Table 1. Simulation Setting I: simulation results from the JM and RM when data are simulated from the JM.

	JM				RM			
	BIAS	SE	SD	CP	BIAS	SE	SD	CP
$\beta_0^{(1)} = 2$	0.006	0.093	0.100	0.980	0.023	0.094	0.101	0.970
$\beta_1^{(1)} = 1$	–0.001	0.111	0.104	0.945	–0.004	0.111	0.104	0.945
$\beta_2^{(1)} = -0.5$	–0.010	0.083	0.079	0.930	0.211	0.075	0.071	0.190
$\beta_3^{(1)} = 0.4$	0.017	0.106	0.100	0.920	–0.073	0.102	0.095	0.880
$\beta_0^{(2)} = 1.8$	–0.012	0.220	0.214	0.965	–0.020	0.212	0.219	0.950
$\beta_1^{(2)} = 0.5$	–0.011	0.123	0.138	0.955	–0.030	0.126	0.140	0.945
$\beta_2^{(2)} = -0.3$	0.016	0.091	0.091	0.930	0.237	0.070	0.076	0.100
$\beta_3^{(2)} = 1$	–0.016	0.116	0.111	0.945	–0.101	0.105	0.103	0.830
$\sigma_1 = 1$	0.005	0.047	0.055	0.970	–0.000	0.050	0.055	0.960
$\sigma_2 = 1$	0.006	0.053	0.055	0.950	–0.061	0.045	0.048	0.760
$\sigma_3 = 1$	0.005	0.053	0.060	0.970	0.001	0.052	0.059	0.980
$\sigma_4 = 1$	0.006	0.059	0.061	0.960	–0.073	0.049	0.051	0.700
$\rho_{12} = 0.4$	–0.006	0.053	0.058	0.950	–0.097	0.058	0.063	0.630
$\rho_{13} = 0.2$	–0.025	0.114	0.116	0.940	–0.047	0.110	0.119	0.930
$\rho_{14} = 0.2$	–0.016	0.064	0.067	0.975	–0.133	0.063	0.067	0.495
$\rho_{23} = 0.1$	–0.012	0.064	0.069	0.955	–0.129	0.064	0.067	0.515
$\rho_{24} = 0.2$	–0.025	0.097	0.105	0.965	–0.165	0.094	0.108	0.665
$\rho_{34} = 0.5$	–0.015	0.063	0.060	0.915	–0.099	0.067	0.064	0.605
$\sigma_e^{(1)} = 0.5$	0.009	0.031	0.032	0.930	0.010	0.032	0.032	0.920
$\sigma_e^{(2)} = 0.5$	0.005	0.037	0.040	0.925	0.008	0.036	0.041	0.930
$\gamma = -1$	–0.032	0.240	0.233	0.950	0.433	0.142	0.128	0.110
$h_0 = 0.1$	0.002	0.017	0.017	0.950	0.033	0.012	0.011	0.125
$\nu_1^{(1)} = -0.8$	–0.040	0.155	0.164	0.960				
$\nu_2^{(1)} = -0.8$	0.003	0.151	0.160	0.955				
$\nu_1^{(2)} = -0.8$	–0.042	0.156	0.158	0.945				
$\nu_2^{(2)} = -0.8$	0.006	0.158	0.158	0.955				

In Setting I, we simulate data from the joint model by assuming $\nu_1^{(1)} = \nu_2^{(1)} = \nu_1^{(2)} = \nu_2^{(2)} = -0.8$ so that the hazard of the terminal event depends on the multidimensional latent variables. Table 1 displays the bias (the average of the posterior means minus the true values), standard error (SE, the square root of the average of the posterior variance), standard deviation (SD, the standard deviation of the posterior means), and coverage probabilities (CP) of 95% equal-tail credible intervals from both the JM (true model) and RM. The results suggest that the joint model provides parameter estimates with very small biases, SE being close to SD, and the coverage probabilities being close to the nominal level of 0.95. In contrast, when all components in $\boldsymbol{\nu}$ are nonzero as in Setting I, incorrectly constraining all these parameters to 0 in the reduced model leads to inaccurate estimation results in the regression parameters of the latent process (i.e. $\boldsymbol{\beta}$) and the parameters in the covariance matrix. For example, the reduced model gives biased estimates and poor CP for the regression parameters $\beta_2^{(1)}, \beta_2^{(2)}$, as well as the variance/covariance parameters σ_2, σ_4, and $\boldsymbol{\rho}$. Moreover, model selection criterion DIC_3 from JM is substantially lower than that from RM in all simulated datasets.

Geskus[30] indicates that in the random effects selection model (the survival process depends on the longitudinal process via random effects, similar to the JM in the current context), the parameters from the longitudinal subprocess can be estimated with little or no bias using only the longitudinal model if either (1) the occurrence of the terminal event only depends on observed longitudinal measurements, or (2) there are many measurements

Table 2. Simulation Setting II: simulation results from the JM and RM when data are simulated from the RM.

	JM				RM			
	BIAS	SE	SD	CP	BIAS	SE	SD	CP
$\beta_0^{(1)} = 2$	0.005	0.088	0.096	0.975	0.007	0.089	0.096	0.965
$\beta_1^{(1)} = 1$	–0.005	0.107	0.102	0.955	–0.007	0.110	0.103	0.955
$\beta_2^{(1)} = -0.5$	–0.009	0.077	0.071	0.935	–0.010	0.073	0.069	0.935
$\beta_3^{(1)} = 0.4$	0.013	0.098	0.094	0.930	0.011	0.098	0.094	0.940
$\beta_0^{(2)} = 1.8$	–0.017	0.210	0.209	0.960	–0.013	0.218	0.212	0.950
$\beta_1^{(2)} = 0.5$	–0.012	0.128	0.137	0.945	–0.010	0.130	0.138	0.950
$\beta_2^{(2)} = -0.3$	0.008	0.092	0.085	0.930	0.001	0.089	0.084	0.945
$\beta_3^{(2)} = 1$	–0.010	0.118	0.106	0.920	–0.004	0.118	0.106	0.935
$\sigma_1 = 1$	0.000	0.048	0.053	0.955	0.002	0.049	0.053	0.950
$\sigma_2 = 1$	0.005	0.048	0.048	0.960	0.007	0.047	0.048	0.965
$\sigma_3 = 1$	0.005	0.053	0.058	0.960	0.008	0.053	0.058	0.950
$\sigma_4 = 1$	0.011	0.051	0.053	0.950	0.013	0.052	0.053	0.950
$\rho_{12} = 0.4$	–0.003	0.052	0.053	0.950	–0.000	0.052	0.053	0.955
$\rho_{13} = 0.2$	–0.026	0.111	0.113	0.945	–0.026	0.113	0.114	0.940
$\rho_{14} = 0.2$	–0.012	0.060	0.061	0.955	–0.012	0.061	0.061	0.955
$\rho_{23} = 0.1$	–0.010	0.059	0.063	0.960	–0.009	0.057	0.063	0.960
$\rho_{24} = 0.2$	–0.015	0.094	0.099	0.955	–0.013	0.096	0.101	0.955
$\rho_{34} = 0.5$	–0.009	0.058	0.056	0.945	–0.010	0.058	0.056	0.960
$\sigma_e^{(1)} = 0.5$	0.004	0.032	0.031	0.935	0.004	0.032	0.031	0.945
$\sigma_e^{(2)} = 0.5$	0.005	0.035	0.039	0.945	0.002	0.038	0.041	0.920
$\gamma = -1$	–0.025	0.153	0.160	0.955	–0.007	0.149	0.157	0.960
$h_0 = 0.1$	–0.001	0.009	0.009	0.955	0.001	0.009	0.009	0.965
$\nu_1^{(1)} = 0$	–0.011	0.129	0.124	0.940				
$\nu_2^{(1)} = 0$	0.010	0.130	0.128	0.945				
$\nu_1^{(2)} = 0$	–0.004	0.157	0.144	0.910				
$\nu_2^{(2)} = 0$	0.009	0.153	0.147	0.935				

per individual, such that they contain sufficient information to estimate the trajectory with little bias. Neither of the two conditions is met in our simulation study. First, the occurrence of the terminal event is correlated with the longitudinal process via the shared random effects, which may also depend on unobserved longitudinal measurements. Second, the median number of measurements per individual in the simulation datasets is 5 (range is $1-7$), which may not contain sufficient information of the trajectory, as compared with 14 CD4 measurements (range is $1-59$) or 8 RNA measurements (range is $1-55$) in the examples from Geskus.[30] Thus, the reduced model without considering the association process provides inaccurate estimation. As pointed out by a reviewer, because the simulated censoring time is independent of the survival process, there should be no bias in the estimate of the regression parameter γ in model (4) when treatment is the only covariate. The large difference in the estimates of γ and the baseline hazard h_0 between two models is due to model non-collapsibility, where random effects are left out from the RM model. The estimates of the outcome-specific parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ are presented in Web Tables 2 and 3. Results from Setting I suggest that when the association between the survival process and the latent variable process is ignored, the estimates for $\boldsymbol{\beta}$ and variance/covariance matrix are biased. However, the ignorance of process association does not impact on the estimation of latent variables as well as $\boldsymbol{a}$ and $\boldsymbol{b}$. A similar observation (reduced models have reasonable estimates of $\boldsymbol{a}$ and $\boldsymbol{b}$) has also been made in our prior works.[16,31]

In Setting II, we simulate data from the reduced model by assuming $\nu_1^{(1)} = \nu_2^{(1)} = \nu_1^{(2)} = \nu_2^{(2)} = 0$ so that the hazard of the terminal event does not depend on any of the latent variables, while all other parameters remain the same. The results in Table 2 and Web Tables 4 and 5 suggest that both models generate comparable results, i.e. the bias is negligible, SE is close to SD, and the credible interval coverage probabilities are reasonably close to 95%. Model selection criterion DIC_3 from both models is comparable in all simulated datasets. Under model overparameterization, the estimates of elements in $\boldsymbol{\nu}$ from the joint model are correctly close to zero.

In conclusion, the simulation results suggest that when the time to the terminal event is dependent on the multivariate longitudinal outcomes, model RM gives inaccurate estimation results in the regression parameters of the latent process and the parameters in the covariance matrix. In contrast, when the terminal event and the multivariate longitudinal outcomes are independent, both the models JM and RM give comparable results.

5 Application to the Ceftriaxone Study

In this section, following Wang and Luo,[17] we apply the proposed joint MLTLMM model and the Bayesian inference framework to the motivating Ceftriaxone study. In Wang and Luo,[17] we detected three dimensions in the 10 items of ALSFRS. Specifically, we compare the joint model and the reduced model with either 3 or 1 latent variables (denoted as JM_{3LV}, RM_{3LV}, JM_{1LV} and RM_{1LV}, respectively). We consider the following covariates for all latent variables: the treatment assignment variable trt_i (1 for active, and 0 for placebo), time in year t_{ij}, time and treatment interaction, and forced vital capacity (FVC) at baseline fvc_i. In addition, we consider the treatment assignment and FVC in the proportional hazards model. Hence, we fit the following model:

$$\theta_i^{(p)}(t_{ij}) = \beta_0^{(p)} + \beta_1^{(p)}\text{trt}_i + \beta_2^{(p)}t_{ij} + \beta_3^{(p)}(\text{trt}_i \times t_{ij}) + \beta_4^{(p)}\text{fvc}_i + u_{i0}^{(p)} + u_{i1}^{(p)}t_{ij} + e_i^{(p)}(t_{ij})$$

$$h_i(t) = h_0(t)\exp\left[\gamma_1\text{trt}_i + \gamma_2\text{fvc}_i + \sum_{p=1}^{P}\left(\nu_1^{(p)}u_{i0}^{(p)} + \nu_2^{(p)}u_{i1}^{(p)}\right)\right]$$

where $\boldsymbol{u}_i = \left(u_{i0}^{(1)}, u_{i1}^{(1)}, \ldots, u_{i0}^{(P)}, u_{i1}^{(P)}\right)' \sim N(0, \Sigma)$, with $\Sigma_{rr} = \sigma_r^2$, $\Sigma_{rs} = \rho_{rs}\sigma_r\sigma_s$ and $r, s = 1, \ldots, 2 \times P$. The reduced model can be obtained by constraining the coefficients $\nu_1^{(p)} = \nu_2^{(p)} = 0$ for all p. Using a 4-Core 3.6 GHz Intel E3-1271 processor, the total runtime is 9 h.

Table 3 compares models JM_{3LV}, RM_{3LV}, JM_{1LV}, and RM_{1LV} using Bayesian model selection criteria. Model JM_{3LV} performs significantly better than the other three models with smaller $\bar{D}$, DIC_3, EAIC and EBIC. Moreover, both joint models JM_{3LV} and JM_{1LV} outperform their reduced model counterparts RM_{3LV} and RM_{1LV}, respectively, suggesting that joint modeling is essential and it improves the model fitting. Web Figure 1 displays the Cox-Snell residual plot from JM_{3LV}, indicating an overall good model fit. Thus, we select model JM_{3LV} as the final model. Table 4 presents the factor loading matrix for model JM_{3LV}. Results in Table 4 are similar and consistent with Table 3 in Wang and Luo,[17] i.e. the three latent variables $\theta^{(1)}$, $\theta^{(2)}$, and $\theta^{(3)}$ are the underlying disease severity associated with bulbar, fine motor, and gross motor functions, respectively.

Table 5 and Web Tables 6, 7 and 8 show the posterior mean, standard deviation (SD), and 95% equal-tail credible intervals of parameters estimate from the best fitting model JM_{3LV}. The results indicate that placebo patients have significant ALS progression at the rate of deterioration of –5.511 unit per year $\left(\hat{\beta}_2^{(1)}\right.$, 95% CI:

Table 3. Model comparison statistics for the Ceftriaxone study.

	$\bar{D}$	DIC_3	EAIC	EBIC
JM_{3LV}	**77830.6**	**80736.1**	**78052.6**	**78522.2**
RM_{3LV}	78119.1	81068.0	78329.1	78773.3
JM_{1LV}	110267.5	111387.7	110395.5	110666.2
RM_{1LV}	110425.5	111548.3	110549.5	110811.8

Note: The best fitting model is highlighted in bold. $\bar{D}$: the posterior mean of the deviance; DIC_3: deviance infromation criterion; EAIC: expected Akaike information criterion; EBIC: expected Bayesian information criterion.

Table 4. Factor loading matrix for model JM_{3LV} with three latent variables.

	JM_{3LV}		
	$\theta^{(1)}$	$\theta^{(2)}$	$\theta^{(3)}$
1. Speech	**1**	0	0
2. Salivation	**0.436**	–0.005	–0.051
3. Swallowing	**0.554**	0.030	0.054
4. Handwriting	–0.046	**0.742**	–0.047
5. Cutting food	0	**1**	0
6. Dressing and hygiene	–0.037	**0.673**	0.206
7. Turning in bed	0.021	**0.388**	**0.360**
8. Walking	0	0	**1**
9. Climbing stairs	0.027	0.065	**0.910**
10. Breathing	0.095	0.045	0.104

Bold values indicate the items the latent variables are loaded on.

Table 5. Results of fitting the joint model JM_{3LV} in the Ceftriaxone study.

	Mean	SD	95% CI	
Parameters for the bulbar function				
Int ($\beta_0^{(1)}$)	3.684	0.922	2.001	5.449
Trt ($\beta_1^{(1)}$)	–0.427	0.527	–1.485	0.569
Time ($\beta_2^{(1)}$)	–5.511	0.370	–6.275	–4.835
Time:Trt ($\beta_3^{(1)}$)	0.160	0.403	–0.578	0.998
FVC ($\beta_4^{(1)}$)	2.059	0.246	1.589	2.548
Parameters for the fine motor function				
Int ($\beta_0^{(2)}$)	7.499	0.829	5.841	9.091
Trt ($\beta_1^{(2)}$)	–0.276	0.426	–1.055	0.669
Time ($\beta_2^{(2)}$)	–7.083	0.420	–7.867	–6.277
Time:Trt ($\beta_3^{(2)}$)	0.323	0.421	–0.416	1.182
FVC ($\beta_4^{(2)}$)	0.145	0.221	–0.291	0.572
Parameters for the gross motor function				
Int ($\beta_0^{(3)}$)	9.083	0.900	7.252	10.791
Trt ($\beta_1^{(3)}$)	0.234	0.476	–0.679	1.190
Time ($\beta_2^{(3)}$)	–7.145	0.446	–8.068	–6.294
Time:Trt ($\beta_3^{(3)}$)	0.291	0.456	–0.562	1.210
FVC ($\beta_4^{(3)}$)	0.645	0.204	0.244	1.071
Parameters for survival				
γ_1	–0.019	0.177	–0.367	0.329
γ_2	–0.318	0.083	–0.475	–0.155
$\nu_1^{(1)}$	–0.072	0.016	–0.105	–0.041
$\nu_2^{(1)}$	–0.288	0.043	–0.377	–0.208
$\nu_1^{(2)}$	0.023	0.017	–0.011	0.056
$\nu_2^{(2)}$	0.147	0.050	0.047	0.247
$\nu_1^{(3)}$	–0.041	0.018	–0.078	–0.007
$\nu_2^{(3)}$	–0.135	0.046	–0.226	–0.046

$[-6.275, -4.835]\Big)$ in the bulbar function, –7.083 unit per year $\left(\hat{\beta}_2^{(2)}, 95\%\text{ CI: } [-7.867, -6.277]\right)$ in the fine motor function, and –7.145 unit per year $\left(\hat{\beta}_2^{(3)}, 95\%\text{ CI: } [-8.068, -6.294]\right)$ in the gross motor function. In comparison, ceftriaxone patients have ALS progression at the rate of deterioration of –5.351 unit per year $\left(\hat{\beta}_2^{(1)} + \hat{\beta}_3^{(1)}, 95\%\text{ CI:}\right.$

$[-6.853, -3.837]\Big)$ with insignificant ceftriaxone treatment effect of slowing down the ALS progression rate by 0.160 per year $\left(\hat{\beta}_3^{(1)}, 95\%\ \text{CI: } [-0.578, 0.998]\right)$ in the bulbar function. Similar interpretations can be made for the other two latent variables.

We observe that $\nu_1^{(1)}$ and $\nu_1^{(3)}$ are negative and significantly different from zero ($\nu_1^{(1)} = -0.072$, 95% CI: $[-0.105, -0.041]$, and $\nu_1^{(3)} = -0.041$, 95% CI: $[-0.078, -0.007]$), suggesting that the ALS patients with worse baseline bulbar and gross motor functions tend to have shorter survival. Moreover, the negative and statistically significant $\nu_2^{(1)}$ and $\nu_2^{(3)}$ ($\nu_2^{(1)} = -0.288$, 95% CI: $[-0.377, -0.208]$, and $\nu_2^{(3)} = -0.135$, 95% CI: $[-0.226, -0.046]$) indicate that the ALS patients with faster progression rates in bulbar and gross motor functions tend to have higher hazard of death. This new interesting finding can only be detected by using our joint MLTLMM model.

6 Discussion

Clinical studies of many complex diseases often witness that the longitudinal outcomes are subject to dependent terminal events. Ignoring such dependence leads to biased estimates. In this article, we have proposed a joint modeling framework consisting of a multivariate latent trait linear mixed model and a proportional hazards model to jointly analyze the multivariate longitudinal data subject to dependent terminal events. Two models are linked together via shared random effects representing the subject-specific disease severity. The proposed joint model has a better fit than the reduced model in the analysis of the Ceftriaxone study. We have found that ALS patients with worse baseline bulbar and gross motor functions and with faster progression rates in these functions tend to have higher death hazard. The simulation studies have shown that in the presence of a dependent terminal event, the joint model successfully recovers the true parameters whereas the reduced model gives biased estimates for the regression parameters and the variance/covariance parameters. Under the scenario of independent terminal event, the joint model provides results comparable with the true reduced model.

There are some limitations in our proposed model that we will address in the near future. First, the multivariate longitudinal outcomes and time to death are linked by shared random effects in this study. This time-independent formulation leads to a closed-form solution for the integral in the survival model and renders great computational advantage. Other functional forms of joint models that allow a reasonable summary of the whole longitudinal trajectories have been proposed.[32] To this end, Rizopoulos et al.[33] investigated various association structures between the longitudinal and event time responses and used Bayesian model averaging (BMA) to combine various joint models with different association structures. However, functional forms in the context of multiple latent variables deserve more research effort. Second, our model for ordinal outcomes requires the proportional odds assumption. Statistical tests to evaluate this assumption in the traditional ordinal logistic regression have been criticized for having a tendency to reject the null hypothesis (that the sets of coefficients are the same between each pair of outcome groups), in cases where the assumption does hold.[34] Tests of the proportional odds assumption in the longitudinal latent variable setting are not well established, and the consequence of violating the assumption is unclear and worth future examination. In addition, we have chosen a multivariate normal distribution for the random effects vector because it is flexible in modeling the covariance structure within and between various types of longitudinal outcomes and it has meaningful interpretation on correlation. In generalized linear mixed models, misspecification of random effects distribution has little impact on the parameters that are not associated with the random effects.[35–37] The impact of random effects misspecification in the proposed modeling framework warrants further investigation. We will also investigate the effect of random effects misspecification and relax the normality assumption by considering Bayesian non-parametric (BNP) framework based on Dirichlet process mixture.[38]

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Sheng Luo's research was supported in part by the National Institute of Neurological Disorders and Stroke (grant no. R01NS091307). The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at

Austin for providing high-performing computing resources that have contributed to the research results reported within this article. URL: http://www.tacc.utexas.edu.

ORCID iD

Sheng Luo http://orcid.org/0000-0003-4214-5809

Supplemental Material

Supplemental material for this article is available online.

References

1. Wijesekera L and Leigh P. Amyotrophic lateral sclerosis. *Orphanet J Rare Dis* 2009; **4**: 3.
2. Berry J, Shefner J, Conwit R, et al. Design and initial results of a multi-phase randomized trial of ceftriaxone in amyotrophic lateral sclerosis. *PLoS One* 2013; **8**: e61177.
3. Prize4Life. Pooled Resource Open-Access ALS Clinical Trials Database; 2013. Data Sets, https://nctu.partners.org/ProACT (accessed 12 March 2018).
4. Abe K, Itoyama Y, Sobue G, et al. Confirmatory double-blind, parallel-group, placebo-controlled study of efficacy and safety of edaravone (MCI-186) in amyotrophic lateral sclerosis patients. *Amyotroph Lateral Scler Frontotemporal Degener* 2014; **15**: 610–617.
5. Bacci E, Staniewska D, Coyne K, et al. Item response theory analysis of the amyotrophic lateral sclerosis functional rating scale-revised in the pooled resource open-access ALS clinical trials database. *Amyotroph Lateral Scler Frontotemporal Degener* 2016; **17**: 157–167.
6. The ACTS Study Group. The amyotrophic lateral sclerosis functional rating scale. Assessment of activities of daily living in patients with amyotrophic lateral sclerosis. *Arch Neurol* 1996; **53**: 141–147.
7. Clarke S, Hickey A, O'Boyle C, et al. Assessing individual quality of life in amyotrophic lateral sclerosis. *Qual Life Res* 2001; **10**: 149–158.
8. Coco D, Marchese S, La Bella V, et al. The amyotrophic lateral sclerosis functional rating scale predicts survival time in amyotrophic lateral sclerosis patients on invasive mechanical ventilation. *Chest* 2007; **132**: 64–69.
9. Bedlack R, Vaughan T, Wicks P, et al. How common are ALS plateaus and reversals? *Neurology* 2016; **86**: 808–812.
10. Balsis S, Unger A, Benge J, et al. Gaining precision on the Alzheimer's disease assessment scale-cognitive: a comparison of item response theory-based scores and total scores. *Alzheimer's Dementia* 2012; **8**: 288–294.
11. Grimby G, Tennant A and Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med* 2012; **44**: 97–98.
12. Andrich D. *Rasch models for measurement*. Newbury Park, USA: Sage, 1988.
13. Samejima F. Estimation of latent ability using a response pattern of graded scores. *ETS Res Bull Series* 1968; **1968**: i–169.
14. Fox J and Glas C. Bayesian estimation of a multilevel IRT model using Gibbs sampling. *Psychometrika* 2001; **66**: 271–288.
15. Glas C, Geerlings H, van de Laar M, et al. Analysis of longitudinal randomized clinical trials using item response models. *Contemp Clin Trials* 2009; **30**: 158–170.
16. Luo S and Wang J. Bayesian hierarchical model for multiple repeated measures and survival data: an application to Parkinson's disease. *Stat Med* 2014; **33**: 4279–4291.
17. Wang J and Luo S. Multidimensional latent trait linear mixed model: An application in clinical studies with multivariate longitudinal outcomes. *Stat Med* 2017; **36**: 3244–3256.
18. Henderson R, Diggle P and Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics* 2000; **1**: 465–480.
19. Faucett CL and Thomas DC. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Stat Med* 1996; **15**: 1663–1685.
20. Wulfsohn MS and Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997; **53**: 330–339.
21. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* 2011; **67**: 819–829.
22. Wang J, Luo S and Li L. Dynamic prediction for multiple repeated measures and event time data: an application to Parkinson's disease. *Ann Appl Stat* 2017; **11**: 1787.
23. Lawless J and Zhan M. Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. *Can J Stat* 1998; **26**: 549–565.
24. Feng S, Wolfe R and Port F. Frailty survival model analysis of the National Deceased Donor Kidney Transplant Dataset using Poisson variance structures. *J Am Stat Assoc* 2005; **100**: 728–735.

25. Liu L and Huang X. Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *J Royal Stat Soc: Ser C* 2009; **58**: 65–81.
26. Stan Development Team. Stan modeling language: user's guide and reference manual, version 2.17.0, 2017, http://mc-stan.org/ (accessed 14 July 2017).
27. Gelman A, Carlin J, Stern H, et al. *Bayesian data analysis*. Boca Raton, USA: CRC Press, 2013.
28. Celeux G, Forbes F, Robert C, et al. Deviance information criteria for missing data models. *Bayesian Analys* 2006; **1**: 651–673.
29. Carlin B and Louis T. *Bayesian methods for data analysis*. Boca Raton, USA: CRC Press, 2009.
30. Geskus RB. Which individuals make dropout informative? *Stat Meth Med Res* 2014; **23**(1): 91–106.
31. He B and Luo S. Joint modeling of multivariate longitudinal measurements and survival data with applications to Parkinson's disease. *Stat Meth Med Res* 2016; **25**: 1346–1358.
32. Brown ER. Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS. *Ann Appl Stat* 2009; **3**: 1163–1182.
33. Rizopoulos D, Hatfield LA, Carlin BP, et al. Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. *J Am Stat Assoc* 2014; **109**: 1385–1397.
34. Harrell F. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. New York, USA: Springer, 2015.
35. Jacqmin-Gadda H, Sibillot S, Proust C, et al. Robustness of the linear mixed model to misspecified error distribution. *Computat Stat Data Analys* 2007; **51**: 5142–5154.
36. Rizopoulos D, Verbeke G and Molenberghs G. Shared parameter models under random effects misspecification. *Biometrika* 2008; **95**: 63–74.
37. McCulloch CE, Neuhaus JM, et al. Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter. *Stat Sci* 2011; **26**: 388–402.
38. Escobar M. Estimating normal means with a Dirichlet process prior. *J Am Stat Assoc* 1994; **89**: 268–277.