Original Research Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
2024, Vol. 33(10) 1783–1799

DOI: 10.1177/09622802241269010
journals.sagepub.com/home/smm
Sage

Flexible joint model for time-to-event and non-Gaussian longitudinal outcomes

Hortense Doms[1], Philippe Lambert[1,2] and Catherine Legrand[1]

Abstract

In medical studies, repeated measurements of biomarkers and time-to-event data are often collected during the follow-up period. To assess the association between these two outcomes, joint models are frequently considered. The most common approach uses a linear mixed model for the longitudinal part and a proportional hazard model for the survival part. The latter assumes a linear relationship between the survival covariates and the log hazard. In this work, we propose an extension allowing the inclusion of nonlinear covariate effects in the survival model using Bayesian penalized B-splines. Our model is valid for non-Gaussian longitudinal responses since we use a generalized linear mixed model for the longitudinal process. A simulation study shows that our method gives good statistical performance and highlights the importance of taking into account the possible nonlinear effects of certain survival covariates. Data from patients with a first progression of glioblastoma are analysed to illustrate the method.

Keywords

Joint models, Bayesian P-splines, longitudinal outcome, survival outcome, generalized linear mixed models

1 Introduction

In medical studies for which the primary interest is to record the time at which a particular event occurs, information on multiple biomarkers is also often collected longitudinally throughout the follow-up period. This provides a combination of survival and longitudinal information about each individual under study. Many methods allow the longitudinal and time-to-event responses to be studied separately, but it is often preferable to analyse them jointly.[1] Indeed, longitudinal measurements could have a predictive role in the analysis of patient survival. A more suitable objective is therefore to measure the association between longitudinal measurements and the event risk. In addition, it is important to take into account that the longitudinal response is typically measured with error and has missing data.

To meet this need, a class of statistical models has been developed: joint models for longitudinal and time-to-event data.[2] The basic idea behind the joint model approach consists in: first, define a model for the longitudinal biomarker measurements; second, define a model for the time-to-event data; and third, link the two submodels with a shared latent association structure. We consider the survival outcome as the primary interest and we incorporate a summary measure of the longitudinal process as a covariate in the time-to-event model.

The seminal formulation of this model was introduced by Faucett and Thomas[2] and Wulfsohn and Tsiatis[3] and is based on a single longitudinal outcome assumed to be normally distributed and a single time-to-event outcome that are linked by adding the expected values of the longitudinal response as a covariate in the event risk model. The joint model framework

[1]Institut de Statistique, Biostatistique et Sciences Actuarielles, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[2]Institut de Mathématiques, Université de Liège, Liège, Belgium

Corresponding author:
Hortense Doms, Institut de Statistique, Biostatistique et Sciences Actuarielles, Université catholique de Louvain, Voie du Roman Pays 20, Louvain-la-Neuve 1348, Belgium.
Email: hortense.doms@uclouvain.be

has subsequently been widely studied in the literature and several extensions have been proposed. These include, among others, non-Gaussian longitudinal outcome,[4–6] multiple longitudinal outcomes[7–9] and multiple recurrent or competing events.[10–12] Chi and Ibrahim[8] presented a joint model considering multiple longitudinal and multiple survival outcomes. Complete overviews of joint modelling techniques have been published.[1,13–15] Another important extension is to allow greater flexibility in joint modelling. Indeed, real datasets often exhibit nonlinear trajectories or effects. Rizopoulos and Ghosh[9] proposed a spline-based approach to describe subject-specific nonlinear longitudinal evolutions. Andrinopoulou et al.[16] described a Bayesian joint model that allows a time-varying coefficient to link the longitudinal and survival submodels, using P-splines.[17] Similarly, Köhler et al.[18] proposed a flexible additive joint model for a continuous longitudinal outcome and single time-to-event outcome under the Bayesian framework using P-splines. This model is highly flexible as it is able to represent all the components of the joint model in an additive and flexible way. However, this approach can only accommodate a Gaussian longitudinal marker.

In this article, we focus on flexible specifications of the survival part in the joint model and more specifically on the potential nonlinearity of the effect of the baseline survival covariates. To our knowledge, no work has focused on assessing whether accounting for nonlinear effects of survival covariates could improve the fit of a joint model while considering a non-Gaussian longitudinal response. We assume nonlinear effects by using Bayesian P-splines.[19] We also use P-splines to describe the logarithm of the baseline hazard. To force smoothness of the fitted curves and to avoid overfitting, a roughness penalty is applied. Our method is valid for Gaussian and non-Gaussian distributions of the longitudinal marker.

The study that motivated our research concerns patients with a first progression of glioblastoma. Glioblastoma is the most common brain cancer in adults, but its survival prognosis is very poor.[20–22] Therefore, it is essential to be able to identify prognostic factors for this cancer in order to guide treatment strategies. We aim to use our methodology on these data to assess the effects of potential prognostic factors more accurately.

The rest of this article is organized as follows. In Section 2, we describe the formulation of the joint model and introduce our proposed flexible survival submodel using P-splines. In Section 3, we present the Bayesian estimation of the model. In Section 4, we discuss a simulation study to evaluate the statistical performance of our approach. Section 5 illustrates our approach to glioblastoma data. Finally, in Section 6, we conclude the article with a discussion.

2 The joint model

2.1 Longitudinal submodel

For every individual ($i = 1, \ldots, n$), we collect longitudinal response values $y_{ij} = y_i(t_{ij})$ observed at time t_{ij} ($j = 1, \ldots, n_i$). The subject-specific evolution over time of the longitudinal response is modelled using a generalized mixed effect model (GLMM).[23] In particular, we postulate

$$\begin{cases} \eta_i(t) = g\{\mathbb{E}(y_i(t) \mid b_i)\} = \boldsymbol{x_i}^T(t)\,\boldsymbol{\beta} + \boldsymbol{z_i}^T(t)\,\boldsymbol{b_i} \\ \boldsymbol{b_i} \sim N(\boldsymbol{0}, \boldsymbol{D}) \end{cases} \tag{1}$$

where $\boldsymbol{\beta}$ denotes the p-dimensional vector of fixed effects associated with the time-dependent design vector $\boldsymbol{x_i}(t)$ and $\boldsymbol{b_i}$ denotes the q-dimensional vector of random effects associated with $\boldsymbol{z_i}(t)$. The random effects $\boldsymbol{b_i}$ allow each of the subjects to have a different response profile in time, and we assume a multivariate normal distribution with zero mean and variance-covariance matrix $\boldsymbol{D}$. Conditional on the random effects $\boldsymbol{b_i}$, the distribution of $\boldsymbol{y_i}$ is assumed to be a member of the exponential family, with $g(.)$ denoting a known link function.[23] When the observed measurements on individuals derive from a counting process, a common choice is to use a logarithm link $g(.) = \log(.)$, which is the canonical link function for the Poisson distribution. When the observed measurements have a binomial distribution, the logit can be used as the canonical link function. In the simple case of normally distributed longitudinal data, the associated canonical link function is the identity function.

2.2 Survival submodel

Let T_i^* denote the true event time for the i-th individual, C_i the right-censoring time, $\delta_i = I(T_i^* \leq C_i)$ the event indicator (1 if the ith individual experiences the event, 0 otherwise) and $T_i = \min(T_i^*, C_i)$ the reported time value. To link the longitudinal and survival outcomes, a common approach in the joint modelling framework assumes a proportional hazards model for the conditional hazard:

$$h_i(t|\mathcal{M}_i(t), \boldsymbol{w_i}) = h_0(t) \exp\left(\boldsymbol{\gamma_w}^T \boldsymbol{w_i} + m\{\mathcal{M}_i(t), b_i, \alpha\}\right), \quad t > 0 \tag{2}$$

where $h_0(t)$ denotes the baseline hazard function and $\boldsymbol{w_i} = [w_{i1}, \ldots, w_{iL}]^T$ is the vector of baseline survival covariates for subject i with regression coefficients $\boldsymbol{\gamma_w}$. In the joint modelling framework, $h_0(t)$ should not be left completely unspecified.[1] Finally, let $\mathcal{N}_i(t) = \{\eta_i(s), 0 \leq s < t\}$ denote the history of the underlying unobserved longitudinal process up to time t. Function $m\{\mathcal{N}_i(t), b_i, \alpha\}$ specifies the association structure between the two submodels with parameter α tuning the strength of the association between the longitudinal process and the survival outcome. Various specifications are used in the literature.[1,14] In this article, we consider the current-value association structure defined by

$$m\{\mathcal{N}_i(t), b_i, \alpha\} = \alpha\eta_i(t) \tag{3}$$

This parameterization is very intuitive and easy to interpret with $\exp(\alpha)$ giving the relative risk increase of an event at time t resulting from a simultaneous one unit increase in $\eta_i(t)$.

2.3 Flexible survival submodel

Model (2) assumes that all survival covariates have a linear effect on the log-hazard. We propose an extended version of this model that allows for possible nonlinear effects of some survival covariates. Specifically, we have

$$h_i(t|\mathcal{N}_i(t), \boldsymbol{w_i}, \boldsymbol{v_i}) = h_0(t)\exp\left\{\boldsymbol{\gamma_w^T w_i} + \sum_{j=1}^{J} f_j(v_{ij}) + \alpha\eta_i(t)\right\}, \quad t > 0 \tag{4}$$

where covariates $\boldsymbol{v_i} = [v_{i1}, \ldots, v_{iJ}]^T$ enter the log-hazard in a nonlinear way and $\boldsymbol{w_i} = [w_{i1}, \ldots, w_{iL}]^T$ is a vector of categorical or additional continuous covariates with a corresponding vector of regression coefficients $\boldsymbol{\gamma_w^T} = [\gamma_{w0}, \ldots, \gamma_{wL}]$. The additive terms $\boldsymbol{f} = [f_1, \ldots, f_J]^T$ are assumed to be smooth and represent nonlinear effects of the associated continuous covariates. The corresponding survival function is defined by

$$S_i(t|\mathcal{N}_i(t), \boldsymbol{w_i}, \boldsymbol{v_i}) = \exp\left(-\int_0^t h_0(s)\exp\left\{\boldsymbol{\gamma_w^T w_i} + \sum_{j=1}^{J} f_j(v_{ij}) + \alpha\eta_i(s)\right\} ds\right) \tag{5}$$

A well-known approach used for modelling and estimating smooth functions relies on B-splines. A B-spline of degree d consists of $d + 1$ polynomial pieces, each of degree d, that join together at d specific points called inner knots.[17] The additive terms f_j are expressed as follows[24]:

$$f_j(v_{ij}) = \sum_{k=1}^{K} \gamma_{v,jk} B_{jk}(v_{ij})$$

where $\gamma_{v,jk}$ represents the B-spline coefficient associated to the B-spline basis function $B_{jk}(.)$ at the respective covariate value v_{ij}. For every f_j, we assume the same number K of basis functions $B_{jk}(.)$ for simplicity (associated to equidistant knots on the range of the jth covariate) and we define the $n \times K$ matrix $\boldsymbol{B_j}$ for which the element at the ith row and kth column correspond to $B_{jk}(v_{ij})$.

As introduced in Section 2.2 and suggested by Lambert and Eilers[25] and by Rizopoulos[1] in the specific context of joint models, we use the same approach to model $h_0(t)$ and express the logarithm of the baseline hazard function as

$$\log\{h_0(t)\} = \sum_{u=1}^{U} \gamma_{h_0,u} B_u(t)$$

where $\gamma_{h_0,u}$ is the regression coefficient associated to the uth function of a B-spline basis associated to equidistant knots on $(0, \max_i T_i)$.

Choosing the optimal number and positions of knots is a complex task that has an impact on the estimated functions. Indeed, as the number of knots increases, the approximation becomes more flexible. However, too many knots can cause serious over-fitting. To overcome this problem, Eilers and Marx[17] proposed a penalized B-splines approach, called P-splines. They suggest to use a relatively large number of equally spaced knots and to counterbalance the flexibility of the fit by adding a roughness penalty based on the differences of adjacent B-spline coefficients.

P-splines can also be estimated in a Bayesian framework as Bayesian P-splines.[19,26] The difference penalties are replaced by their stochastic analogues, that is, random walk priors. We obtain the roughness penalty by imposing a Gaussian random field prior on the spline coefficients[27] (see Section 3.2 for details). Both the log baseline hazard function and the additive terms $\boldsymbol{f}$ will be approximated using Bayesian P-splines.

3 Bayesian estimation

We apply a Bayesian approach where inference is based on the posterior of the model parameters. In particular, we estimate the parameters of the joint models (1) and (4) using Markov chain Monte Carlo (MCMC) methods.

3.1 Likelihood function

In the framework of joint models for longitudinal and survival data, the likelihood is derived under the common assumption that the longitudinal and survival processes are independent conditionally on the random effects $\boldsymbol{b_i}$. Moreover, the longitudinal measurements $\boldsymbol{y_i}$ of each individual are assumed independent given the random effects. Formally we have,[6]

$$\begin{aligned} p(T_i, \delta_i, \boldsymbol{y_i}|\boldsymbol{b_i}; \boldsymbol{\theta}) &= p(T_i, \delta_i|\boldsymbol{b_i}; \boldsymbol{\theta})\, p(\boldsymbol{y_i}|\boldsymbol{b_i}; \boldsymbol{\theta}) \\ p(\boldsymbol{y_i}|\boldsymbol{b_i}; \boldsymbol{\theta}) &= \prod_{j=1}^{n_i} p(y_i(t_{ij})|\boldsymbol{b_i}; \boldsymbol{\theta}) \end{aligned} \tag{6}$$

where $\boldsymbol{\theta} = [\boldsymbol{\theta_s}^T, \boldsymbol{\theta_y}^T, \boldsymbol{\theta_b}^T]^T$ contains $\boldsymbol{\theta_s}$ the parameters for the survival outcome, $\boldsymbol{\theta_y}$ are the parameters for the longitudinal outcome and $\boldsymbol{\theta_b}$ are the parameters of the random-effect covariance matrix. Function $p(.)$ generically denotes the corresponding probability or density function.

The expression of the marginal likelihood contribution for the ith subject is written as

$$\begin{aligned} L_i(\boldsymbol{\theta}; T_i, \delta_i, \boldsymbol{y_i}) &= \int p(T_i, \delta_i, \boldsymbol{y_i}, \boldsymbol{b_i}; \boldsymbol{\theta})\, db_i \\ &= \int p(T_i, \delta_i|\boldsymbol{b_i}; \boldsymbol{\theta_s}, \boldsymbol{\beta}) \left[\prod_{j=1}^{n_i} p\left(y_i(t_{ij})|\, \boldsymbol{b_i}; \boldsymbol{\theta_y}\right)\right] p(\boldsymbol{b_i}; \boldsymbol{\theta_b})\, db_i \end{aligned}$$

The marginal log-likelihood function formulated for all subjects is obtained by $l(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log L_i(\boldsymbol{\theta}; T_i, \delta_i, \boldsymbol{y_i})$.

The extended expression for the marginal likelihood contribution of the ith subject in the survival submodel is given by

$$\begin{aligned} p(T_i, \delta_i|\boldsymbol{b_i}; \boldsymbol{\theta_s}, \boldsymbol{\beta}) &= \left[h_i(T_i|\mathcal{N}_i(T_i); \boldsymbol{\theta_s}, \boldsymbol{\beta})\right]^{\delta_i} S_i(T_i|\mathcal{N}_i(T_i); \boldsymbol{\theta_s}, \boldsymbol{\beta}) \\ &= \left[h_0(T_i) \exp\left\{\boldsymbol{\gamma_w}^T \boldsymbol{w}_i + \sum_{j=1}^{J} f_j(v_{ij}) + \alpha \eta_i(T_i)\right\}\right]^{\delta_i} \\ &\quad \times \exp\left(-\int_0^{T_i} h_0(s) \exp\left\{\boldsymbol{\gamma_w}^T \boldsymbol{w}_i + \sum_{j=1}^{J} f_j(v_{ij}) + \alpha \eta_i(s)\right\} ds\right) \\ &= \exp\left[\sum_{u=1}^{U} \gamma_{h_0,u} B_u(T_i) + \boldsymbol{\gamma_w}^T \boldsymbol{w}_i + \sum_{j=1}^{J}\sum_{k=1}^{K} \gamma_{v,jk} B_{jk}(v_{ij}) + \alpha \eta_i(T_i)\right]^{\delta_i} \\ &\quad \times \exp\left[-\exp\left\{\boldsymbol{\gamma_w}^T \boldsymbol{w}_i + \sum_{j=1}^{J}\sum_{k=1}^{K} \gamma_{v,jk} B_{jk}(v_{ij})\right\} \int_0^{T_i} \exp\left\{\sum_{u=1}^{U} \gamma_{h_0,u} B_u(s) + \alpha \eta_i(s)\right\} ds\right] \end{aligned} \tag{7}$$

The distribution of the outcomes $\boldsymbol{y_i}$ is assumed to belong to the exponential family of distributions and so the likelihood contribution of the longitudinal submodel takes the form

$$p(y_{ij}|\boldsymbol{b_i}; \boldsymbol{\theta_y}) = \exp\left\{\frac{y_{ij}\, \psi_{ij}(\boldsymbol{b_i}) - c\{\psi_{ij}(\boldsymbol{b_i})\}}{a(\phi)} - d(y_{ij}, \phi)\right\} \tag{8}$$

where $\psi_{ij}(\boldsymbol{b_i})$ and ϕ are the natural and scale parameters, respectively, and $c(.)$ and $d(.)$ are known functions that identify the member of the exponential family. For one-parameter distributions, functions $a(\phi)$ and $d(y_{ij}, \phi)$ can be simplified to a and $d(y_{ij})$, respectively.[23] Different choices of these functions to cover the most common distributions are presented in Table A in the Appendix.

The normality assumption for the distribution of random effects implies that

$$p(\boldsymbol{b_i}; \boldsymbol{\theta_b}) = (2\pi)^{-q/2} \det(\boldsymbol{D})^{-1/2} \exp(-\boldsymbol{b_i^t}\, \boldsymbol{D}^{-1} \boldsymbol{b_i}/2) \tag{9}$$

where q denotes the dimension of the random effect vector.

A numerical method must be used for the evaluation of the integral in (7) because it has no closed-form solution. We approximate it using the 15-point Gauss-Kronrod quadrature formula.

3.2 Priors and posterior distributions

Bayesian estimation of model parameters is based on the posterior distribution. The expression for the posterior distribution of the model parameters is derived under the assumptions defined in (6). Using Bayes theorem, we obtain

$$p(\boldsymbol{\theta}, \boldsymbol{b} \mid \boldsymbol{T}, \boldsymbol{\delta}, \boldsymbol{y}) \propto \prod_{i=1}^{n} \prod_{j=1}^{n_i} p(T_i, \delta_i \mid \boldsymbol{b_i}; \boldsymbol{\theta_s})\, p(y_{ij} \mid \boldsymbol{b_i}; \boldsymbol{\theta_y})\, p(\boldsymbol{b_i}; \boldsymbol{\theta_b})\, p(\boldsymbol{\theta}) \tag{10}$$

where $\boldsymbol{\theta}$ denotes the full parameter vector with joint prior $p(\boldsymbol{\theta})$. The expressions for $p(T_i, \delta_i \mid \boldsymbol{b_i}; \boldsymbol{\theta})$, $p(y_{ij} \mid \boldsymbol{b_i}; \boldsymbol{\theta})$, and $p(\boldsymbol{b_i}; \boldsymbol{\theta_b})$ can be found in (7), (8), and (9), respectively.

We use standard non-informative prior distribution for $\boldsymbol{\theta}$.[14,28] In particular, independent univariate diffuse normal priors are used for $\boldsymbol{\beta}$ (the vector of fixed effects in the longitudinal submodel), for $\boldsymbol{\gamma_w}$ (the vector of regression coefficients in the survival submodel) and for α (the association parameter). For the standard deviations of the variance-covariance matrix of the random effects $\boldsymbol{D}$, we take a Gamma prior.[29] A random sample from the joint posterior in (10) is obtained using a Metropolis-within-Gibbs algorithm with Gibbs steps for the variance parameter and Metropolis steps for the other parameters.

As already mentioned, the roughness penalty from the frequentist P-spline approach[17] takes the form of a Gaussian Markov field prior for the B-spline parameters in the Bayesian framework,[27]

$$\begin{aligned} p(\boldsymbol{\gamma_v} \mid \tau_v) &\propto \tau_v^{\rho(\boldsymbol{K})/2} \exp\left(-\frac{\tau_v}{2}\, \boldsymbol{\gamma_v}^T \boldsymbol{K}\, \boldsymbol{\gamma_v}\right) \text{ and } \quad \tau_v \sim \text{Gamma}(a_v, b_v) \\ p(\boldsymbol{\gamma_{h_0}} \mid \tau_{h_0}) &\propto \tau_{h_0}^{\rho(\boldsymbol{K_0})/2} \exp\left(-\frac{\tau_{h_0}}{2}\, \boldsymbol{\gamma_{h_0}}^T \boldsymbol{K_0}\, \boldsymbol{\gamma_{h_0}}\right) \text{ and } \quad \tau_{h_0} \sim \text{Gamma}(a_{h_0}, b_{h_0}) \end{aligned} \tag{11}$$

where τ_v and τ_{h_0} are roughness penalty parameters and $\boldsymbol{K} = \boldsymbol{\Delta_r}^T \boldsymbol{\Delta_r}$ is the penalty matrix of rank $\rho(\boldsymbol{K})$ associated to the additive terms f_j, where $\boldsymbol{\Delta}_r$ denotes the rth-order difference penalty matrix (with $r = 2$, say). Similarly, $\boldsymbol{K_0}$ denotes the penalty matrix associated to the baseline hazard function. The amount of smoothness is controlled by τ_v and τ_{h_0}. It is generally recommended to set a_v and a_{h_0} equal to 1 and b_v and b_{h_0} equal to a small number (say 0.005).[19] Alternatively, a mixture of Gamma distributions could be considered to obtain a robust specification for the roughness penalty prior.[26]

3.3 Identifiability

Given the additive structure of the model, we are forced to impose constraints to obtain an identifiable model. We use the sum-to-zero constraint described by Wood,[30] such that

$$\sum_{i=1}^{n} f_j(v_{ij}) = 0$$

for $j = 1, \ldots, J$. The sum of the values of each additive function, at the observed covariate values, must be zero. To achieve this condition, the column mean can be subtracted from each column of the matrix $\boldsymbol{B}$, leading to the centred B-spline matrix $\tilde{\boldsymbol{B}}$. Since the rank of the matrix $\tilde{\boldsymbol{B}}$ reduces to $K - 1$, we have to impose an additional constraint. The last element of each spline vector $\boldsymbol{\gamma_{vj}}$ is set to zero and we delete the last column of $\tilde{\boldsymbol{B}}$ and of the difference matrix $\boldsymbol{\Delta_r}$, yielding $\tilde{\boldsymbol{K}}$ for the penalty matrix.

4 Simulations

The objective of this section is to evaluate the statistical performance of our approach (`FlexibleJM`) which allows to estimate in a Bayesian approach a joint model that can contain covariates with nonlinear effects in the survival submodel. We evaluate the finite sample properties of our estimation strategy through a simulation study, focusing on two aspects. First, we assess the ability to estimate smooth additive terms, and second, we compare our results with those obtained with the `JMbayes2`[29] package assuming linearity for the additive terms. This package enables to fit a joint model under a Bayesian approach but currently does not permit the inclusion of a nonlinear effects of baseline covariates in the survival submodel. It assumes a linear relationship between each survival covariate and the log hazard function.[6] With the `JMbayes2` package, the estimation of the joint model is performed via the function `jm()`. We name it below the `LinearJM` approach. The `FlexibleJM` approach was implemented using parts of the `JMbayes2` package code as a starting point. The R code used to implement the `FlexibleJM` approach is available upon request or downloadable from https://github.com/hortensedoms.

4.1 Simulation design

In this article, we consider the special case of a Poisson longitudinal response. We also carried out simulations with other distributions for the longitudinal response, the results of which are presented in Tables B and C and Figures A and B in the Appendix. For every setting, we simulate a maximum of 10 repeated measurements for n subjects at a fixed grid of time points $t_p = [0, 2, 4, \ldots, 20]$ based on a true longitudinal model

$$\eta_i(t) = \beta_0 + \beta_1 t_{ij} + \beta_2 x_{1i} + \beta_3 x_{2i} + b_{0i} + b_{1i} t_{ij}$$

where $\beta_0 = 0.55$, $\beta_1 = -0.25$, $\beta_2 = -0.60$, $\beta_3 = 0.30$, $x_{1i} \sim N(0, 0.5^2)$, $x_{2i} \sim \text{Bin}(1, 0.5)$ and $b_{0i} \sim N(0, 0.25^2)$ and $b_{1i} \sim N(0, 0.1^2)$. We generate the observed longitudinal measurements from $y_{ij}|b_{0i} \sim \text{Pois}(\lambda_{ij})$, where $\eta_{ij} = \log(\lambda_{ij})$ and $\eta_{ij} = \eta_i(t_{ij})$. The hazard $h_i(t)$ for every subject is defined as follows:

$$h_i(t|\mathcal{N}_i(t), w_i, v_i) = h_0(t) \exp\left\{ \gamma_1 w_{1i} + \gamma_2 w_{2i} + \sum_{j=1}^{J} f_j(v_{ij}) + \alpha \eta_i(t) \right\}$$

where $\gamma_1 = 0.90$, $\gamma_2 = 0.70$, $\alpha = -0.60$, $w_{1i} \sim N(0, 0.5^2)$, $w_{2i} \sim \text{Bin}(1, 0.5)$ and using the true baseline hazard function $h_0(t) = 1.5(t + 0.9)^{0.45} - 1.2$. For the additive terms part $\sum_{j=1}^{J} f_j(v_{ij})$, we investigated the two following scenarios.
Scenario I. One nonlinear effect:

$$h_i(t) = h_0(t) \exp\{\gamma_1 w_{1i} + \gamma_2 w_{2i} + f_1(v_{i1}) + \alpha \eta_i(t)\}$$

Scenario II. One nonlinear effect and one linear effect:

$$h_i(t) = h_0(t) \exp\{\gamma_1 w_{1i} + \gamma_2 w_{2i} + f_1(v_{i1}) + f_2(v_{2i}) + \alpha \eta_i(t)\}$$

where in both cases

$$f_1(v_{i1}) = -\sin((v_{i1} + 30)/20) - 4; \quad f_2(v_{i2}) = -0.07 v_{i2} + 1.6$$

and the variable v_{i1} is simulated from a uniform distribution on [18, 89] and v_{i2} is simulated from a uniform distribution on [18, 35].

Based on $h_i(t)$, simulated survival times are generated for every subject using the approach described by Crowther and Lambert.[31] The cumulative hazard is evaluated using numerical quadrature and survival times are generated using an iterative algorithm which uses the numerical root finding method of Brent.[32] Every subject is censored after $\max(t_p)$ and we additionally right-censored the time-to-event data by generating subject-specific right-censoring times from a uniform distribution on $\text{U}(k \times \max(t_p), 1.1 \times \max(t_p))$ to the survival times. The value of k is selected in order to obtain 10% and 30% right-censoring rates (cr).

To ensure convergence, we run the model estimation with 20,000 iterations after an adaptive phase of 3000 iterations and a burn-in of 3000. The chains are thinned by 10 in order to finally keep 2000 iterations for each parameter. Geweke statistics[33] are used as a diagnostic tool for chain convergence. For each regression parameter, we assess bias, the root-mean-squared error (RMSE) and the effective coverage of 95% credible intervals (COV). We also assess the integrated

root-mean-squared error (RMISE) for the estimated additive terms. In summary, we have

$$\text{Bias}_{\hat{\beta}_p} = \frac{1}{S}\sum_{s=1}^{S}\left(\hat{\beta}_p^{(s)} - \beta_p\right)$$

$$\text{RMSE}_{\hat{\beta}_p} = \left\{\frac{1}{S}\sum_{s=1}^{S}\left(\hat{\beta}_p^{(s)} - \beta_p\right)^2\right\}^{\frac{1}{2}}$$

$$\text{COV}_{\hat{\beta}_p} = \frac{1}{S}\sum_{s=1}^{S}\mathbb{1}\left(\beta_p \in \text{CI}_{\beta_p}^{(s)}\right)$$

$$\text{RMISE}_{\hat{f}_j} = \left\{\frac{1}{S}\sum_{s=1}^{S}\int\left(\hat{f}_j^{(s)}(v) - f_j(v)\right)^2 dv\right\}^{\frac{1}{2}}$$

where S is the number of replications, $\hat{\beta}_p$ is the estimated posterior mean of a regression parameter β_p and $\mathbb{1}$ is an indicator function that is equal to one if the estimated 95% credible interval (CI) contains the true parameter value.

The first simulation is based on Scenario I. We aim to assess the impact of the sample size and of the right-censoring rate on the quality of the model estimation and, more specifically, on the estimation of the nonlinear additive term f_1. We therefore consider three data settings: Setting *a* where $n = 500$ and cr = 10%, Setting *b* where $n = 500$ and cr = 30% and Setting *c* where $n = 250$ and cr = 10%. The second simulation is based on Scenario II which contains two additives terms f_1 and f_2. We consider $n = 500$ and cr = 10% and we compare the estimation obtained with the methods `FlexibleJM` and `LinearJM`. We would like to highlight that accounting for the possible nonlinear effects of some survival covariates improves the performance of parameter estimation in the joint model. Simulations are performed using $S = 500$ replicates for Scenarios I and II and we used eight B-spline basis functions to estimate the additive terms.

4.2 Simulation results

Figure 1 shows the estimation of the nonlinear additive function f_1 for the three data settings under Scenario I. For each graph, the 500 grey curves represent the estimations of the corresponding unknown smooth function which is represented by the black curve. For each data setting, the observed estimates are close to the target, and as expected the more information we have, the less variability there is. Indeed, if we compare Setting *a* with Setting *b*, we notice that the estimated curves are closer to the target curve in the first case. This is due to the fact that on average 450 events occur in Setting *a* versus 350 in Setting *b*. We also note that for smaller datasets (Setting *c*), there is more uncertainty in the estimation, especially for values at the extremities of the covariate range. The integrated root mean square error is calculated and compared for each setting. We can see in Table 1 that the RMISE values for Settings *b* and *c* are, respectively, 1.6 times and 1.9 times larger than those for Setting *a*. This confirms our previous observation that the estimates obtained in Setting *a* tend to be more precise than in the other two settings.

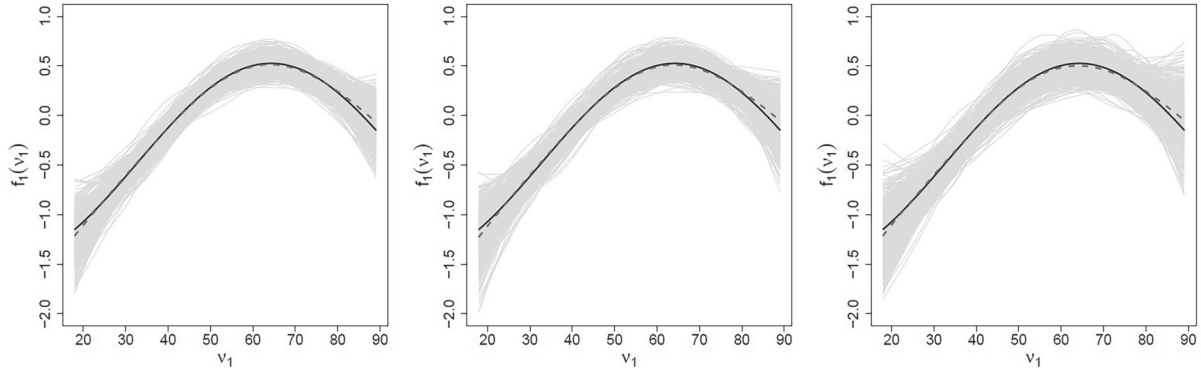

Figure 1. True (black) and estimated (grey) smooth additive term f_1 for $S = 500$ dataset replications in Settings *a, b* and *c* (respectively, left, middle and right) under Scenario I. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves.

Table 1. Simulation results for Scenario I and Scenario II with the `FlexibleJM` and `LinearJM` methods: Integrated root-mean-square error (RMISE) for the smooth additive terms f_1 and f_2. We consider different sample sizes (n) and censoring rates (cr) for Scenario I.

Scenario I				
	n	cr	f	`FlexibleJM`
Setting a :	500	10%	$\hat{f}_1$	0.923
Setting b :	500	30%	$\hat{f}_1$	1.137
Setting c :	250	10%	$\hat{f}_1$	1.778

Scenario II		
f	`FlexibleJM`	`LinearJM`
$\hat{f}_1$	0.930	7.881
$\hat{f}_2$	0.105	0.091

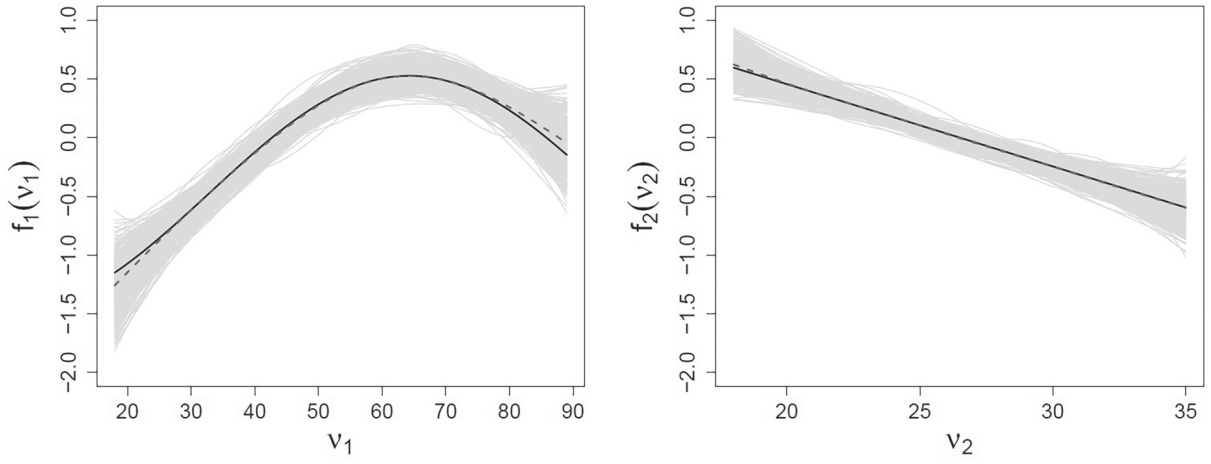

Figure 2. True (black) and estimated (grey) smooth additive terms f_1 and f_2 for $S = 500$ dataset replications in Scenario II obtained with the `FlexibleJM` approach. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves.

The second simulation compares the estimation results obtained with the `FlexibleJM` and `LinearJM` methods within the framework of a joint model allowing the inclusion of nonlinear covariate effects in the survival submodel. Figure 2 and Figure C (in the Appendix) displays the true and the estimated additive terms f_1 and f_2 for $S = 500$ replications obtained respectively with the `FlexibleJM` and the `LinearJM` methods. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves. The average computation times for a dataset are similar with `FlexibleJM` and `LinearJM`, with, respectively, 1.83 and 1.66 min using just one core of a 2.6 GHz Intel Core Processor i5-1145G7.

We notice in Figure 2 that the estimated curves match the true functions (f_1 and f_2) quite well. Also, the dashed lines and the black lines overlap for both functions. It suggests that the estimation of the additive terms using P-splines in the joint model works well using our `FlexibleJM` approach. In Figure C, the 500 grey lines represent the estimated effects of $\boldsymbol{\nu_1}$ and $\boldsymbol{\nu_2}$ obtained using `LinearJM`. As expected, this method cannot take into account the nonlinear effects of survival covariates due to model misspecification, as linear effects are assumed throughout.

Table 2 compares the performance measures obtained for each regression parameter with each of the two methods. While both approaches show similar performances for the linear regression parameters $\beta_0, \ldots, \beta_3$ in the longitudinal submodel, the `FlexibleJM` method outperforms `LinearJM` in estimating the regression parameters γ_1 and γ_2 and the association parameter α, due to the presence of nonlinear effects in the survival part. `FlexibleJM` achieves lower bias and RMSE and better (frequentist) coverages for the credible intervals in the estimation of the survival regression parameters compared to `LinearJM`. This is confirmed by the boxplots of the estimated values in Figure 3. We also note that the RMISE values for $\hat{f}_2$ are quite similar for both methods (see Table 1). This confirms that, compared to the `LinearJM` method, the precision of the estimates provided by `FlexibleJM` is comparable for linear effects and is improved for nonlinear effects.

In conclusion, our simulations show that the `FlexibleJM` approach is able to take into account nonlinear effects of survival covariates and provides accurate estimates of the additives terms. Using `FlexibleJM`, the estimations of the joint model parameters outperforms those obtained with `LinearJM` for the survival parameters.

Table 2. Simulation results for $S = 500$ replications of sample size $n = 500$ in Scenario II with the `FlexibleJM` and `LinearJM` approaches : bias, root mean square error (RMSE) and effective coverage (COV) of 95% credible intervals for the regression parameters.

		`FlexibleJM`			`LinearJM`		
	True	Bias	RMSE	COV	Bias	RMSE	COV
β_0	0.55	0.002	0.071	0.958	0.002	0.071	0.958
β_1	−0.25	−0.003	0.012	0.944	−0.003	0.011	0.938
β_2	−0.60	−0.002	0.046	0.950	−0.002	0.046	0.948
β_3	0.30	0.002	0.089	0.946	0.001	0.089	0.948
γ_1	0.90	0.010	0.069	0.956	−0.045	0.081	0.856
γ_2	0.70	0.003	0.107	0.964	−0.033	0.115	0.918
α	−0.60	0.007	0.072	0.976	0.030	0.079	0.914

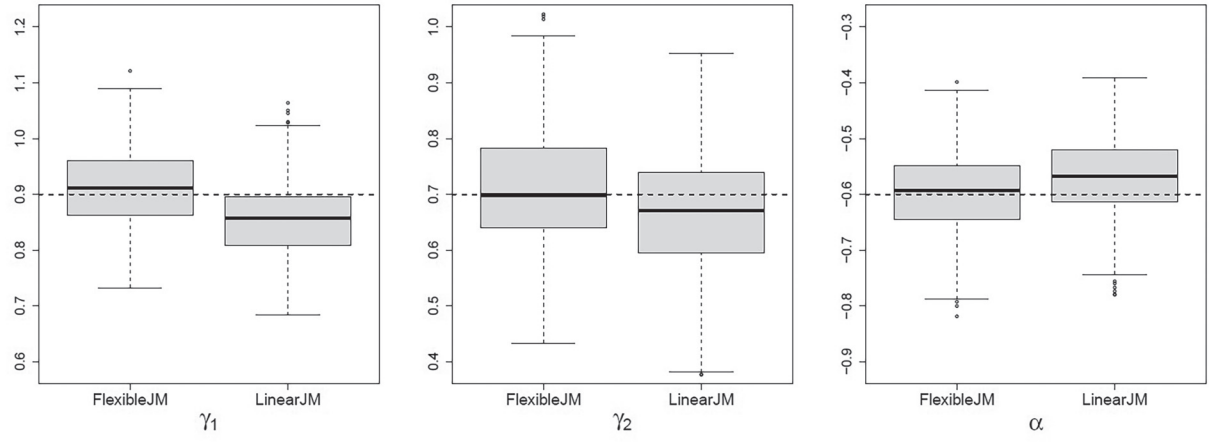

Figure 3. Boxplots of the estimated survival regression parameters from $S = 500$ replications in Scenario II comparing methods `FlexibleJM` and `LinearJM`. The dashed horizontal lines indicate the true value of the parameters.

5 Application

As introduced in Section 1, glioblastoma is a highly malignant form of cancer that develops in the brain. It is a prevalent and extremely aggressive cancer with a very poor prognosis for survival and no curative treatment.[20–22] The assessment of prognostic factors is an important issue to guide treatment decisions and patient management. Previous research has shown that O6-methylguanine-DNA methyltransferase (MGMT) promoter methylation status, corticosteroid administration at baseline, World Health Organization (WHO) performance status (PS), tumour size and, to some extent age have a significant association with patient survival.[34] All of these factors were measured at baseline. Between 2011 and 2015, a phase III clinical trial was conducted by the European Organisation for Research and Treatment of Cancer (EORTC-26101 trial)[35] to determine whether the combined use of lomustine and bevacizumab could enhance overall survival (OS) in patients with first progression of glioblastoma as compared to lomustine alone. In this study, WHO-PS and tumour size were assessed longitudinally throughout the follow-up period. We would therefore like to assess their effect on overall survival over time instead of just accounting for their baseline value at the start of the trial. In addition, corticosteroids are often used to treat patients with glioblastoma to help manage symptoms, but their use must be carefully monitored and managed as their potential side effects can be severe. We therefore want to improve the modelling of the effect of corticosteroids on the risk of death. Given these circumstances, we first want to assess the association between changes over time in WHO-PS and the risk of death in patients with progressive glioblastoma. In a second step, we also want to investigate whether there is an association between changes over time in tumour size and the risk of death. For both analyses, we consider a potential nonlinear effect of the baseline dose of corticosteroids on the risk of death. The dataset analysed contains $n = 516$ patients enrolled across multiple institutions within Europe. Follow-up times varied between 8 and 919 days and 95 (18.4%) patients were right-censored. The median survival time is equal to 267 (95% CI: 249–290). There are 1809 repeated measurements for each biomarker. The mean number of repeated measurements per patient is 3.5 (range: 1–15).

To estimate the joint models presented in Sections 5.1 and 5.2 in the Bayesian framework, MCMC methods are used with chains of length 20,000 following an adaptive phase of 3000 iterations and a burn-in of 3000, see Section 3 for methodological details. The trace plots obtained showed no apparent trend, which can be taken as evidence of convergence. Furthermore, we used the Gelman-Rubin diagnostic to confirm that the scale reduction $\hat{R}$ of all parameters are smaller than 1.1. We also produced a graphical posterior predictive check by comparing the real longitudinal data and the simulated data from the fitted models. In Figure D in the Appendix, we see no systematic discrepancies between the real and simulated data.

5.1 Longitudinal evolution of the WHO performance status

Consider first the evolution over time of the patients' performance status, assessed longitudinally during the follow-up. The WHO Performance Status describes a patient's ability to take care of themselves, as well as their daily activity and physical abilities.[36] This score runs from 0 to 5, with 0 denoting perfect health and 5 denoting death. The interest is to assess longitudinally the patient's level of capacity (good or poor), knowing that a good performance status is typically characterized by a score in 0-2.[37] The performance status is evaluated at baseline and every 6 weeks until week 24, then every 3 months during the follow-up. The binary indicators $y(t)$ of a poor performance status at time t have a Bernoulli distribution with success probability $\pi_i(t)$. We considered the following joint model:

$$\begin{cases} h_i(t) = h_0(t)\exp\{\gamma_1 w_{\texttt{LesionB}} + \gamma_2 w_{\texttt{Age}} + \gamma_3 w_{\texttt{MGMT}_1} + \gamma_4 w_{\texttt{MGMT}_2} + \gamma_5 w_{\texttt{MGMT}_3} + f(\texttt{logDoseC}) + \alpha\eta_i(t)\}, \\ \eta_i(t) = \text{logit}(\pi_i(t)) = \beta_0 + b_{0i} + \beta_1 t \end{cases}$$

In the survival submodel, $f(\texttt{logDoseC})$ is an unknown smooth function estimated in `FlexibleJM`, but forced to be linear in `LinearJM`. Covariate `LesionB` denotes the lesion size at baseline measured by the maximum diameter (in mm) of the largest lesion. Covariate `Age` denotes the age of the patient at baseline. The categorical covariate `MGMT` denotes the methylation status of the MGMT promoter with categories 0 = unmethylated (reference category), 1 = methylated, 2 = undetermined and 3 = not done. This last category refers to patients for whom there was no tumor tissue sample of sufficient quality to be able to carry out the test. Finally, `logDoseC` is constructed using $\texttt{logDoseC} = \log(\texttt{DoseC} + 1)$ where `DoseC` refers to the corticosteroid dose (in mg) administered at baseline. Patients characteristics are described in Table 3.

The estimated parameters of the survival submodel are presented in Table 4. The association between changes in WHO-PS values and the risk of death is significantly positive. Poor performance status are significantly associated with a higher risk of death. The hazard ratio for death with methylated MGMT status as compared with unmethylated MGMT status is $\exp(-0.9015) = 0.41$, which means that the MGMT promoter methylation is an important prognostic factor. The log pseudo-marginal likelihood (LPML) and deviance information criterion (DIC) are better with the `FlexibleJM` approach. Moreover, the left panel of Figure 4 shows that our approach `FlexibleJM` suggests a significant nonlinear effect of the logarithm of the baseline corticosteroid dose `logDoseC` on the risk of death (conditionally on LesionB, Age, MGMT and the current WHO performance status.). The methodology used to assess the significance of a smooth term is described by Wood.[38]

5.2 Longitudinal evolution of lesion size

Consider now the association between the longitudinal evolution of the lesion size and the risk of death. The size of the lesion is assessed for the first time at baseline and every 6 weeks until week 24, then every 3 months.[20] As can be seen from Figure E in the Appendix, part of the patients showed nonlinear evolutions in their lesion size values. To capture the nonlinear subject-specific evolutions, we assume a flexible linear mixed-effects model. We opted for natural cubic splines because the `JMBayes2` package allows the use of the function `ns()` from the package `spline()` in the specification of the longitudinal submodel.[6,29] Therefore, we considered the following joint model:

$$\begin{cases} h_i(t) = h_0(t)\exp\{\gamma_1 w_{\texttt{WHOB}_{\texttt{Poor}}} + \gamma_2 w_{\texttt{Age}} + \gamma_3 w_{\texttt{MGMT}_1} + \gamma_4 w_{\texttt{MGMT}_2} + \gamma_5 w_{\texttt{MGMT}_3} + f(\texttt{logDoseC}) + \alpha\eta_i(t)\}, \\ \eta_i(t) = \beta_0 + s(t) + b_{0i} + b_{1i} t \end{cases}$$

where $s(.)$ is a nonlinear function of t, b_{0i} is a random intercept and $b_{1i}t$ is a random slope.[1,9] The longitudinal measurements of the lesion size $y_i(t)$ are supposed normally distributed conditionally on the random effects, so we have $\eta_i(t) = \mathbb{E}(y_i(t) \mid b_i)$ with the identity link. The binary covariate `WHOB` represents the patient's level of capacity (good or poor) at baseline.

The results in Table 5 show that the association between the time-varying lesion size and the risk of death is significantly positive. We note that the effect of MGMT_2 becomes non-significant when the model is estimated with `FlexibleJM`. As we can see in the right panel of Figure 4, we still get a significant nonlinear effect of `logDoseC` at a baseline.

Table 3. Descriptive statistics of survival covariates and longitudinal biomarkers.

Covariates	
Sample size	516
LesionB	
Median	45.0
Std. Deviation	21.1
Range	11.0–132.0
Age	
Median	57.8
Std. Deviation	10.8
Range	19.3–82.3
logDoseC	
Median	0.4
Std. Deviation	1.5
Range	0.0–5.3
MGMT	
0	169 (32.7%)
1	135 (26.2%)
2	96 (18.6%)
3	116 (22.5%)
WHOB	
Poor	488 (94.6%)
Good	28 (5.4%)
Biomarkers	
Number of repeated measurements	1809
Longitudinal values of lesion	
Median	38.0
Std. Deviation	23.9
Range	0.0–179.0
Longitudinal values of WHO	
Poor	1686 (93.2%)
Good	123 (6.8%)

MGMT: O6-methylguanine-DNA methyltransferase; WHO: World Health Organization.

Table 4. Estimation results of the survival submodel parameters when we consider the binary version of the WHO-PS as the longitudinal biomarker (`WHOL`): estimates and 95% credible intervals are presented. Values of goodness of fit criterion: log pseudo-marginal likelihood (LPML) and deviance information criterion (DIC) are presented.

	`FlexibleJM`		`LinearJM`	
	Est.	95% CI	Est.	95% CI
γ_1 (`LesionB`)	0.3179	[0.1721; 0.4683]	0.3227	[0.1511; 0.5174]
γ_2 (`Age`)	−0.0087	[−0.1772; 0.1600]	−0.0122	[−0.1829; 0.1449]
γ_3 ($\texttt{MGMT}_1$)	−0.9015	[−1.3565; −0.5447]	−0.9646	[−1.5594; −0.4977]
γ_4 ($\texttt{MGMT}_2$)	−0.3690	[−0.8491; 0.1120]	−0.3706	[−0.9193; 0.1116]
γ_5 ($\texttt{MGMT}_3$)	−0.4076	[−0.7578; −0.1043]	−0.4894	[−0.9265; −0.0587]
α ($\texttt{Assoc:WHOL}_{\texttt{Poor}}$)	0.7137	[0.5097; 0.9887]	0.7204	[0.4689; 1.0991]
f (`logDoseC`)	(See Figure 4(a))		0.1345	[0.0296; 0.2466]
LPML	−4388.2		−4418.8	
DIC	8443.4		8445.5	

MGMT: O6-methylguanine-DNA methyltransferase; CI: credible interval.

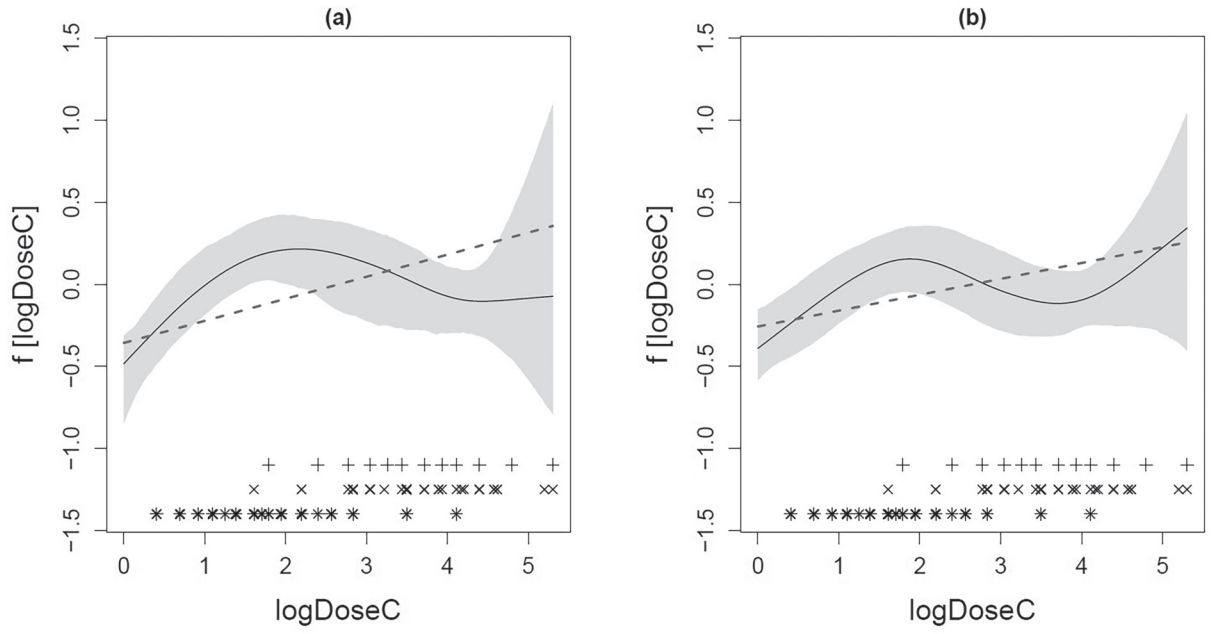

Figure 4. The black curves represent the estimation of the nonlinear effect of `logDoseC` on the log hazard function obtained with the `FlexibleJM` method. The grey surfaces correspond to a 95% pointwise credible envelope. The dotted grey curves represent the effect of `logDoseC` estimated linearly with `LinearJM`. The crosses represent the observed values of `logDoseC`. The left panel (a) corresponds to the joint model involving the longitudinal evolution of a binary version of WHO-PS and the right panel (b) corresponds to the joint model in which we consider the longitudinal evolution of lesion size. The symbols ⋆,× represent the values of `logDoseC` observed for dexamethasone and methylprednisolone, respectively. The category represented by the symbol + includes other types of corticosteroid.

Table 5. Estimation results of the survival submodel parameters when we take the lesion size as longitudinal biomarker (`LesionL`): estimates and 95% credible intervals (CIs) are presented. Values of goodness of fit criterion: log pseudo-marginal likelihood (LPML) and deviance information criterion (DIC) are presented.

	`FlexibleJM`		`LinearJM`	
	Est.	95% CI	Est.	95% CI
γ_1 ($\text{WHOB}_{\text{Poor}}$)	1.2606	[0.7708; 1.7135]	1.2424	[0.7047; 1.7436]
γ_2 (Age)	0.0805	[−0.0224; 0.1821]	0.0704	[−0.0352; 0.1777]
γ_3 (MGMT_1)	−0.5121	[−0.7929; −0.2353]	−0.5374	[−0.8111; −0.2606]
γ_4 (MGMT_2)	−0.1958	[−0.4986; 0.1008]	−0.2306	[−0.5264; −0.0613]
γ_5 (MGMT_3)	−0.1882	[−0.4649; −0.0859]	−0.2486	[−0.5237; −0.0210]
α (Assoc:LesionL)	0.5766	[0.4855; 0.6706]	0.5638	[0.4707; 0.6588]
f (logDoseC)	(See Figure 4(b))		0.0967	[−0.0159; 0.2028]
LPML	−4380.1		−4413.3	
DIC	8743.4		8748.4	

MGMT: O6-methylguanine-DNA methyltransferase.

The estimated nonlinear effects for the log dose of corticosteroid reported in Figure 4 deserve some more comments. Indeed, the two panels in Figure 4 show that an increase in `logDoseC` from 0 to 2 mg is associated to an increased risk of death. Thereafter, the risk seems to decrease, and above a `logDoseC` of 4 mg, the risk stabilizes (left panel) or increases again (right panel) depending on the considered longitudinal biomarker. In order to explain this nonlinear behaviour, we conducted a post-hoc analysis and we found that the different types of corticosteroid administered to patients across centres were not directly comparable. In particular, if we compare the two most commonly administered corticosteroid in this study, a 0.75 mg dose of dexamethasone is equivalent to a 4 mg dose of methylprednisolone.[39] The nonlinear estimation of the effect of `logDoseC`, therefore, highlighted the non-equivalence of the different corticosteroids used, a feature that the `LinearJM` approach would not have revealed.

6 Discussion

We have presented an extension of joint models for longitudinal and survival data that allows us to account for the nonlinearity of the effects of certain covariates in the survival submodel and the possible non-normality of the longitudinal data. These nonlinear effects are modelled using P-splines and implemented in a Bayesian framework with an automatic tuning of the smoothness of the additive terms. Since the longitudinal process is modelled using GLMMs, it allows for the inclusion of a variety of responses variables within the exponential family.

The proposed estimation strategy performs well in different simulation scenarios. Our approach `FlexibleJM` can effectively incorporate nonlinear effects for covariates in the survival submodel and produce accurate estimates of additive terms, thereby extending the `LinearJM` method forcing the user to assume linearity or to categorize the continuous covariate of interest. Taking these nonlinear effects into account also improves the estimation of the conditional effects of the other covariates. Furthermore, we have shown that even when the data are generated with a linear effect for the survival covariate, our approach performs almost as well as the classical linear method in estimating this linear effect. These results suggest that the proposed joint model is worth considering for joint modelling of longitudinal and survival data when a nonlinear effect of a survival covariate is suspected.

One limitation of our method is that the estimated nonlinear effects are common to all patients. `FlexibleJM` does not allow different nonlinear effects to be computed for different groups of patients. Also, it is not recommended to include in the risk model a variable constructed from the associated longitudinal response (e.g. values at time 0). However, it is possible to include the same covariate in both submodels of the joint model.[40]

Within the framework of joint models, several additional extensions are possible. In this article, we have used only one longitudinal outcome, but models that with multiple longitudinal outcomes could be desirable. Different types of longitudinal responses could be accommodated by postulating a multivariate generalized linear mixed effects model. This extension is in theory straightforward as long as we assume that the random effects terms take into account all dependencies between the observed data. That is, given random effects, both the longitudinal outcomes as well as the repeated measurements for each outcome are conditionally independent. Furthermore, the longitudinal outcomes are independent of the time-to-event conditionally on the random effects.[7,9] However, the estimation of multivariate joint models can be very complex since the dimension of the variance-covariance matrix for random effects increases with the number of longitudinal outcomes modelled, resulting in a high-computational burden.[14] As mentioned in Section 2.2, several parameterizations could also be considered, including sophisticated association structures estimated using P-splines.[16] Future research could also explore the use of several ordinal responses in the longitudinal submodel.

Acknowledgements

The authors acknowledge the support of the ARC project IMAL (grant 20/25-107) financed by the Wallonia-Brussels Federation and granted by the Académie universitaire Louvain. The authors thank the European Organization for Research and Treatment of Cancer for permission to use the data from EORTC-26101. The contents of this publication and methods used are solely the responsibility of the authors and do not necessarily represent the official views of the EORTC.

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.

ORCID iDs

Hortense Doms https://orcid.org/0009-0002-4670-2641
Philippe Lambert https://orcid.org/0000-0002-3670-3328

References

1. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data, With Applications in R*. Boca Raton: Chapman & Hall, 2012.
2. Faucett C and Thomas D. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Stat Med* 1996; **15**: 1663–1685.
3. Wulfsohn M and Tsiatis A. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997; **53**: 330–339.
4. Faucett CL, Schenker N and Elashoff RM. Analysis of censored survival data with intermittently observed time-dependent binary covariates. *J Am Stat Assoc* 1998; **93**: 427–437.
5. Viviani S, Alfo M and Rizopoulos D. Generalized linear mixed joint model for longitudinal and survival outcomes. *Stat Comput* 2014; **24**: 417–427.

6. Rizopoulos D. The R package JMbayes for fitting joint models for longitudinal and time-to-event data using MCMC. *J Stat Softw* 2016; **72**: 1–45.
7. Brown ER, Ibrahim JG and DeGruttola V. A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics* 2005; **61**: 64–73.
8. Chi Y and Ibrahim JG. Joint models for multivariate longitudinal and multivariate survival data. *Biometrics* 2006; **62**: 432–445.
9. Rizopoulos D and Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. *Stat Med* 2011; **30**: 1366–1380.
10. Elashoff RM, Li G and Lin N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics* 2008; **64**: 762–771.
11. Williamson P, Kolamunnage-Dona R, Philipson P, et al. Joint modelling of longitudinal and competing risks data. *Stat Med* 2008; **27**: 6426–6438.
12. Andrinopoulou ER, Rizopoulos D, Takkenberg JJM, et al. Joint modeling of two longitudinal outcomes and competing risk data. *Stat Med* 2014; **33**: 3167–3178.
13. Tsiatis AA and Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Stat Sin* 2004; **14**: 809–834.
14. Papageorgiou G, Mauff K, Tomer A, et al. An overview of joint modeling of time-to-event and longitudinal outcomes. *Annu Rev Stat Appl* 2019; **6**: 223–240.
15. Wu L, Liu W, Yi GY, et al. Analysis of longitudinal and survival data: joint modeling, inference methods, and issues. *J Probab Stat* 2012; **2012**: 1–17.
16. Andrinopoulou ER, Rizopoulos D, Takkenberg JJM, et al. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Stat Methods Med Res* 2017; **26**: 1787–1801.
17. Eilers PHC and Marx BD. Flexible smoothing with B-splines and penalties. *Stat Sci* 1996; **11**: 89–121.
18. Köhler M, Umlauf N, Beyerlein A, et al. Flexible Bayesian additive joint models with an application to type 1 diabetes research. *Biometrical J* 2017; **59**: 1144–1165.
19. Lang S and Brezger A. Bayesian P-splines. *J Comput Graph Stat* 2004; **13**: 183–212.
20. Wick W, Gorlia T, Bendszus M et al. Lomustine and bevacizumab in progressive glioblastoma. *New England J Med* 2017; **377**: 1954–1963.
21. Taal W, Oosterkamp HM, Walenkamp AM et al. Combination of bevacizumab and lomustine in patients with first recurrence of glioblastoma: prolongs progression-free survival but not overall survival: a randomized, phase III trial of the EORTC brain tumor and radiotherapy groups. *J Clin Oncol* 2014; **32**: 1239–1247.
22. Tan AC, Ashley DM, L'opez GY, et al. Management of glioblastoma: state of the art and future directions. *CA: A Cancer J Clinici* 2020; **70**: 299–312.
23. Brown H and Prescott R. *Applied Mixed Models in Medicine*. 2nd ed. New York: John Wiley and Sons, 2006.
24. Marx BD and Eilers PHC. Direct generalized additive modeling with penalized likelihood. *Comput Stat Data Anal* 1998; **28**: 193–209.
25. Lambert P and Eilers PHC. Bayesian proportional hazards model with time varying regression coefficients: a penalized Poisson regression approach. *Stat Med* 2005; **24**: 3977–3989.
26. Jullion A and Lambert P. Robust specification of the roughness penalty prior distribution in spatially adaptive Bayesian P-splines models. *Comput Stat Data Anal* 2007; **51**: 2542–2558.
27. Rue H and Held L. *Gaussian Markov Random Fields: Theory and Applications*. Boca Raton: CRC Press, 2005.
28. Alsefri M, Sudell M, Garcia-Finana M, et al. Bayesian joint modelling of longitudinal and time to event data: a methodological review. *BMC Med Res Methodol* 2020; **20**: 1–17.
29. Rizopoulos D, Papageorgiou G and Afonso P. *JMbayes2: Extended Joint Models for Longitudinal and Time-to-Event Data*, 2023. URL http://CRAN.R-project.org/package=JMbayes2, R package version 0.4-5.
30. Wood S. *Generalized Additive Models: An Introduction With R*. Boca Raton, London, New York: Chapman and Hall/CRC Press, 2017.
31. Crowther MJ and Lambert PC. Simulating biologically plausible complex survival data. *Stat Med* 2013; **32**: 4118–4134.
32. Brent RP. Some efficient algorithms for solving systems of nonlinear equations. *SIAM J Numer Anal* 1973; **10**: 327–344.
33. Geweke J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. *Bayesian Stat* 1992; **4**: 641–649.
34. Gorlia T, Stupp R, Brandes AA et al. New prognostic factors and calculators for outcome prediction in patients with recurrent glioblastoma: a pooled analysis of EORTC brain tumour group phase I and II clinical trials. *Eur J Cancer* 2012; **48**: 1176–1184.
35. ClinicalTrials.gov: National Library of Medicine (US). Identifier NCT01290939, Bevacizumab and Lomustine for Recurrent GBM, https://clinicaltrials.gov/ct2/show/NCT01290939 (2021, accessed 27 April 2023).
36. ECOG-ACRIN Cancer Research Group. ECOG Performance Status, https://ecog-acrin.org/resources/ecog-performance-status/ (2022, accessed 27 April 2023).
37. Chahal M, Thiessen B and Mariano C. Treatment of older adult patients with glioblastoma: moving towards the inclusion of a comprehensive geriatric assessment for guiding management. *Curr Oncol* 2022; **29**: 360–376.
38. Wood SN. On P-values for smooth components of an extended generalized additive model. *Biometrika* 2013; **100**: 221–228.
39. Singer M and Webb AR. *Oxford Handbook of Critical Care*. Oxford: Oxford University Press, 2005.
40. Ibrahim JG, Chu H and Chen LM. Basic concepts and methods for joint models of longitudinal and survival data. *J Clini Oncol: Off J Am Soc Clini Oncol* 2010; **28**: 2796–2801.

Appendix

Table A. Expression for a(.), c(.) and d(.) for member of exponential family distribution.

Distribution	a	$c(\psi)$	$d(y)$
Bernoulli	1	$\log(1+\exp(\psi))$	1
Binomial	$1/n$	$\log(1+\exp(\psi))$	$\log[n!/((ny)!\,(n-ny)!)]$
Poisson	1	$\exp(\psi)$	$-\log(y!)$
	$a(\phi)$	$c(\psi)$	$d(y, a(\phi))$
Normal	ϕ	$\psi^2/2$	$-1/2\{y^2/\phi + \log(2\pi\phi^2)$

Table B. Gaussian longitudinal response : $y_{ij}|b_{0i} \sim \mathrm{N}(\lambda_{ij}, \sigma^2)$; $\eta_{ij} = \lambda_{ij}$.

		`FlexibleJM`			`LinearJM`		
	True	Bias	RMSE	COV	Bias	RMSE	COV
β_0	0.55	0.000	0.056	0.958	0.000	0.056	0.958
β_1	−0.25	0.000	0.005	0.954	0.000	0.006	0.954
β_2	−0.60	0.000	0.033	0.958	0.001	0.033	0.960
β_3	0.30	0.000	0.068	0.948	0.000	0.068	0.948
γ_1	0.90	0.003	0.063	0.954	−0.055	0.082	0.788
γ_2	0.70	0.011	0.098	0.964	−0.041	0.107	0.916
α	−0.60	0.006	0.050	0.976	0.042	0.064	0.864

Simulation results for $S = 500$ replications of sample size $n = 500$ in Scenario II with the `FlexibleJM` and `LinearJM` approaches : bias, root mean square error (RMSE) and effective coverage (COV) of 95% credible intervals for the regression parameters.

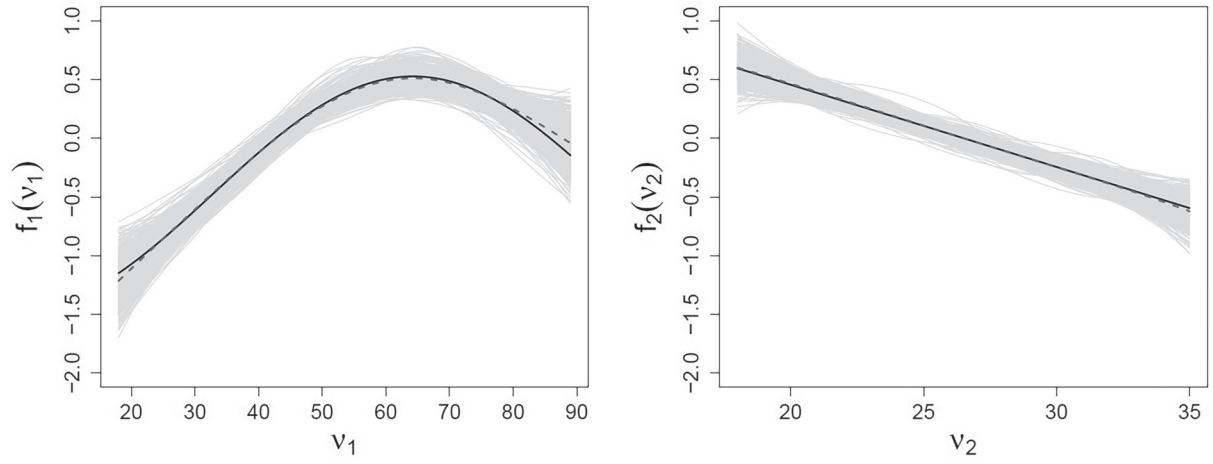

Figure A. Gaussian longitudinal response: $y_{ij}|b_{0i} \sim \mathrm{N}(\lambda_{ij}, \sigma^2)$; $\eta_{ij} = \lambda_{ij}$. True (black) and estimated (grey) smooth additive terms f_1 and f_2 for $S = 500$ dataset replications in Scenario II obtained with the `LinearJM` approach. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves.

Table C. Binary longitudinal response: $y_{ij}|b_{0i} \sim \text{Bernoulli}(\lambda_{ij})$; $\eta_{ij} = \text{logit}(\lambda_{ij})$.

		`FlexibleJM`			`LinearJM`		
	True	Bias	RMSE	COV	Bias	RMSE	COV
β_0	0.55	0.008	0.105	0.934	−0.009	0.104	0.951
β_1	−0.25	−0.002	0.019	0.929	−0.003	0.030	0.922
β_2	−0.60	−0.004	0.064	0.959	0.003	0.066	0.949
β_3	0.30	0.000	0.115	0.944	0.000	0.117	0.939
γ_1	0.90	0.010	0.078	0.952	−0.037	0.081	0.886
γ_2	0.70	0.012	0.112	0.966	−0.035	0.114	0.931
α	−0.60	−0.005	0.091	0.972	0.019	0.100	0.921

Simulation results for $S = 500$ replications of sample size $n = 500$ in Scenario II with the `FlexibleJM` and `LinearJM` approaches : bias, root mean square error (RMSE) and effective coverage (COV) of 95% credible intervals for the regression parameters.

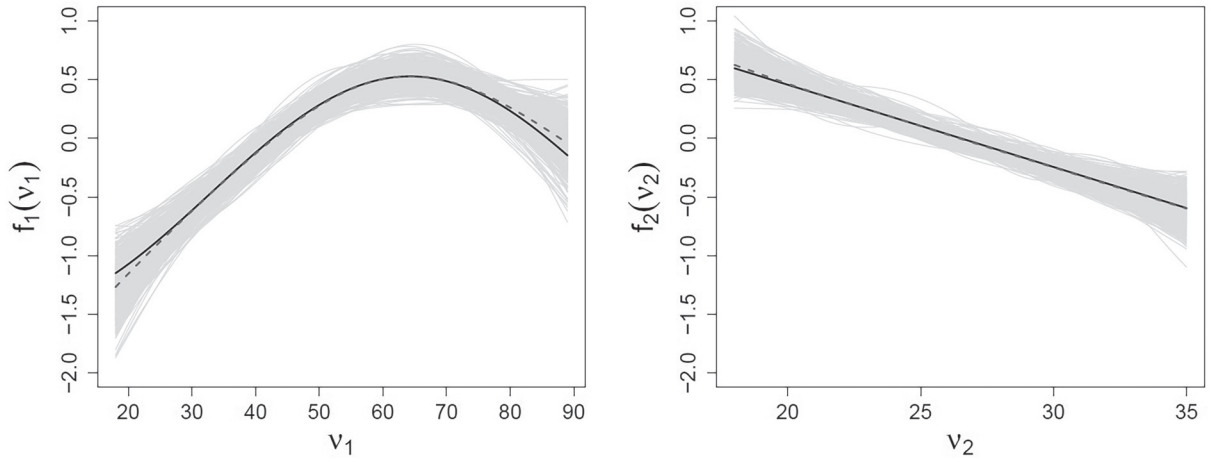

Figure B. Binary longitudinal response: $y_{ij}|b_{0i} \sim \text{Bernoulli}(\lambda_{ij})$; $\eta_{ij} = \text{logit}(\lambda_{ij})$. True (black) and estimated (grey) smooth additive terms f_1 and f_2 for $S = 500$ dataset replications in Scenario II obtained with the `LinearJM` approach. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves.

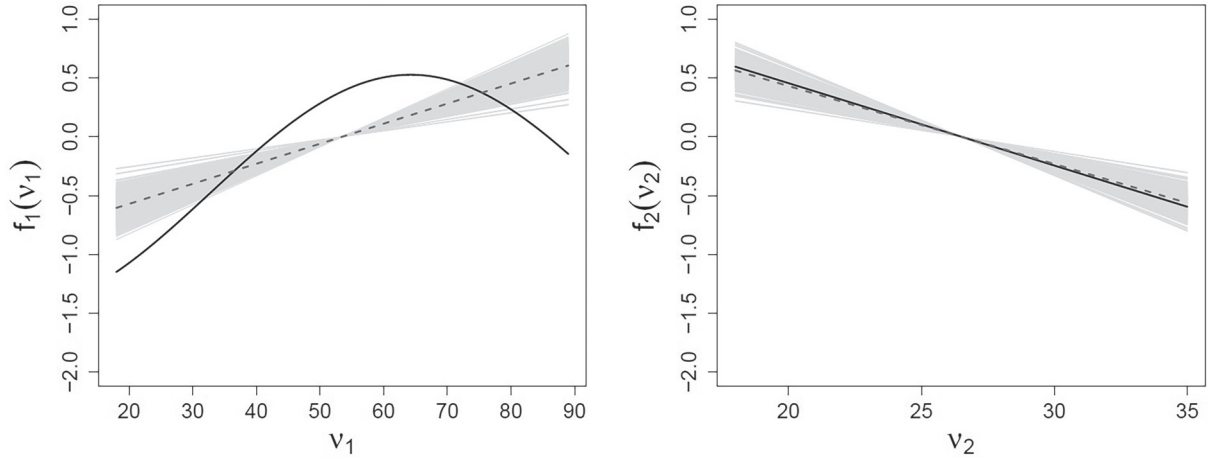

Figure C. True (black) and estimated (grey) smooth additive terms f_1 and f_2 for $S = 500$ dataset replications in Scenario II obtained with the `LinearJM` approach. The grey dashed curves correspond to the pointwise mean of the 500 estimated curves.

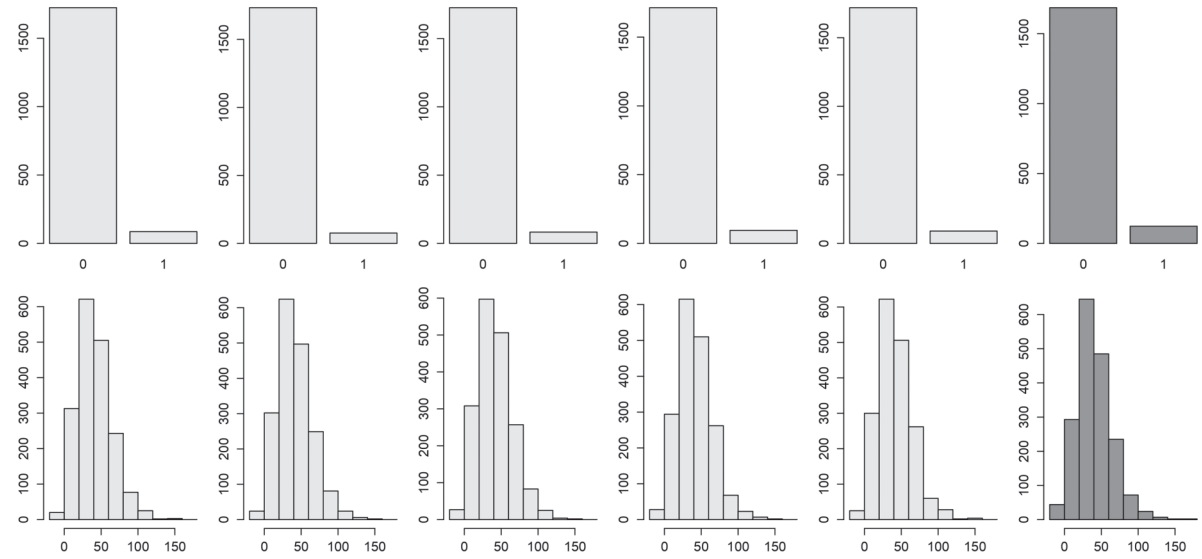

Figure D. Histograms of five replications (light grey) and true values (dark grey) of both longitudinal responses from the posterior distributions.

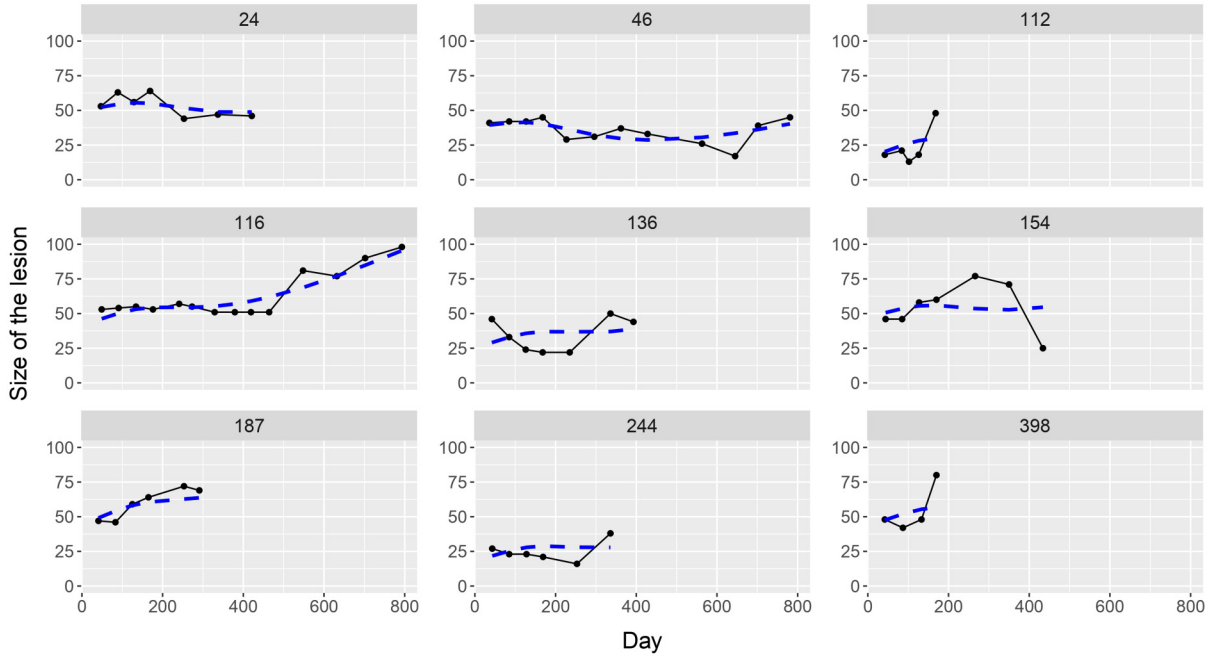

Figure E. Subject-specific longitudinal profiles (black) and model predictions (blue) of lesions size for nine randomly selected patients.