Bayesian Influence Measures for Joint Models for Longitudinal and Survival Data

Hongtu Zhu[1,*], **Joseph G. Ibrahim**[1,**], **Yueh-Yun Chi**[2,***], and **Niansheng Tang**[3,****]

[1]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7420, U.S.A

[2]Department of Biostatistics, University of Florida, Gainesville, Florida 32611-7450, U.S.A

[3]Department of Statistics, Yunnan University, Kunming, Yunnan Province, P.R. of China

Summary

This article develops a variety of influence measures for carrying out perturbation (or sensitivity) analysis to joint models of longitudinal and survival data (JMLS) in Bayesian analysis. A perturbation model is introduced to characterize individual and global perturbations to the three components of a Bayesian model, including the data points, the prior distribution, and the sampling distribution. Local influence measures are proposed to quantify the degree of these perturbations to the JMLS. The proposed methods allow the detection of outliers or influential observations and the assessment of the sensitivity of inferences to various unverifiable assumptions on the Bayesian analysis of JMLS. Simulation studies and a real data set are used to highlight the broad spectrum of applications for our Bayesian influence methods.

Keywords

Bayesian influence measure; Longitudinal; Perturbation model; Sensitivity analysis; Survival

1. Introduction

There has been extensive research literature on joint modeling of longitudinal and survival data (JMLS) by using either frequentist or Bayesian methods. For instance, the early development of joint models for longitudinal and survival data was primarily motivated by characterizing the relationship between features of CD4 or viral load profiles and time-to-event in HIV/AIDS clinical trials. JMLS has been further developed in other types of biomedical applications, such as cancer vaccine (immunotherepy) trials and quality of life studies. References include Pawitan and Self (1993); De Gruttola and Tu (1994); Tsiatis, DeGruttola, and Wulfsohn (1995); Faucett and Thomas (1996); Wulfsohn and Tsiatis (1997); Henderson, Diggle, and Dobson (2000); Wang and Taylor (2001); Xu and Zeger (2001); Law, Taylor, and Sandler (2002); Song, Davidian, and Tsiatis (2002); Chen, Ibrahim, and Sinha (2002, 2004a); Brown and Ibrahim (2003a; Brown and Ibrahim (2003b;

Brown, Ibrahim, and DeGruttola (2005) and Chi and Ibrahim (2006, 2007) among many others. Nice overviews of JMLS are given in Tsiatis and Davidian (2004) and Yu et al. (2004) from a frequentist perspective, and in Ibrahim, Chen, and Sinha (2001) as well as Hanson, Branscum, and Johnson (2011) from a Bayesian perspective.

Recent advances in computation and prior elicitation have made Bayesian analysis of these "complex" JMLS feasible. For instance, some Bayesian semiparametric approaches for longitudinal profiles include Gaussian processes, and functional Dirichlet processes among others. Nonparametric prior processes for the baseline cumulative hazard function, include the gamma process prior, the correlated gamma process, and the Dirichlet process prior among others. However, very little has been done on developing a general Bayesian influence approach to detect influential points and to assess various unverifiable assumptions underlying the JMLS, which is the focus of this article.

Bayesian local and global influence (or robustness) approaches have been widely used to perturb the data, the prior and the sampling distribution and assess the influence of these perturbations on the posterior distribution and the associated posterior quantities. However, these Bayesian local and global influence approaches are not directly applicable to complex JMLS. Although there are some frequentist diagnostic tools (Dobson and Henderson, 2003; Rizopoulos, Verbeke, and Molenberghs, 2008; Rizopoulos and Ghosh, 2011) developed specifically for some JMLS, they are not sufficient for carrying out sensitivity analysis (e.g., prior) of complex Bayesian analysis of JMLS.

The development of the proposed methodology was primarily motivated by a clinical trial conducted by the International Breast Cancer Study Group (IBCSG) (Chi and Ibrahim, 2006). A subset of the IBCSG data set contains the longitudinal measurements for quality of life (QOL) measured at baseline and at months 3 and 18 after randomization. In addition, we have bivariate failure times, including disease-free survival (DFS) and overall survival (OS), which were collected from $n = 832$ patients from Switzerland, Sweden, and New Zealand/ Australia. The covariates are different therapeutic procedures, age, estrogen receptor (ER) status (negative/positive), and the number of positive nodes of the tumor. Although a Bayesian analysis of joint longitudinal and bivariate survival models have been used to fit this data set (Chi and Ibrahim, 2006), a general diagnostic framework for assessing such a model fit to the IBCSG data is completely lacking. There is a great need to develop various diagnostic measures for the detection of outliers and/or influential observations, and the assessment of the sensitivity of inferences to the prior distributions and other unverifiable assumptions on the JMLS.

The article is organized as follows. In Section 2, we introduce a general model for jointly modeling multivariate longitudinal and survival data. In Section 3, we discuss various perturbation models and then calculate their associated Bayesian influence measures to quantify the effects of perturbing the data, the prior, and the sampling distribution on possible posterior quantities of interest. We present a detailed analysis of the IBCSG data in Section 4.

2. Joint Models of Longitudinal and Survival Data

Consider data from n independent subjects. For each subject, we observe a $K \times 1$ vector of multiple longitudinal responses and an $M \times 1$ vector of multivariate time-to-event outcomes. For the ith subject, let $y_{ik}(t_{ijk})$ be an assessment of the kth longitudinal response measured at time t_{ijk} and let $\mathbf{Y}_{ik} = (y_{ik}(t_{i1k}), \ldots, y_{ik}(t_{in_{ik}k}))^T$ denote the observed longitudinal process for the kth response for $i = 1, \ldots, n, j = 1, \ldots, n_{ik}$, and $k = 1, \ldots, K$. Moreover, for the ith

subject, we observe the event time $T_{im} = \min(T^*_{im}, C_{im})$ and the event indicator $\delta_{im} = \mathbf{1}(T^*_{im} \leq C_{im})$

for the mth time-to-event outcome for $m = 1, \ldots, M$, where $\mathbf{1}(A)$ is an indicator function of an event A and T^*_{im} and C_{im} denote the true event time and the censoring time, respectively.

We consider a general shared parameter model for jointly modeling the longitudinal and survival data as follows. Let $\boldsymbol{b}_i = (\boldsymbol{b}_{i1}, \ldots, \boldsymbol{b}_{iK})$ be time-independent random effects underlying both the longitudinal and survival processes for the ith subject. Conditional on $\boldsymbol{b}_i$, all components of the longitudinal outcomes and the time-to-event outcomes are independent. Let $\mathbf{Y}_i = (\mathbf{Y}_{i1}, \ldots, \mathbf{Y}_{iK})$, $\mathbf{T}_i = (T_{i1}, \ldots, T_{iM})^T$, and $\delta_i = (\delta_{i1}, \ldots, \delta_{iM})^T$. The shared parameter model is defined by

$$p(\mathbf{Y}_i, \mathbf{T}_i, \delta_i; \theta) = \int p(\mathbf{Y}_i | \boldsymbol{b}_i; \theta_y) p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \theta_T) p(\boldsymbol{b}_i; \theta_b) d\boldsymbol{b}_i, \quad (1)$$

where $p(\ldots | \ldots)$ and $p(\ldots)$ denote the appropriate conditional density and density functions, respectively, and $\theta^T = (\theta^T_y, \theta^T_T, \theta^T_b)$ is the vector containing all unknown parameters corresponding to each of the submodels. Moreover, corresponding to the partition of $\theta^T = (\theta^T_y, \theta^T_T, \theta^T_b)$, $\theta^T = (\theta^T_F, \theta^T_I)$ can be further decomposed as a finite-dimensional parameter vector $\theta^T_F = (\theta^T_{y,F}, \theta^T_{T,F}, \theta^T_{b,F})$ and a vector of infinite-dimensional parameters $\theta^T_I = (\theta^T_{y,I}, \theta^T_{T,I}, \theta^T_{b,I})$, such as the baseline hazard function or cumulative baseline hazard function, for each of the submodels. This class of shared parameter models in equation (1) includes most JMLS in the existing literature as a special case (Ibrahim et al., 2001; Tsiatis and Davidian, 2004; Hanson et al., 2011).

We specify each of the submodels in equation (1) as follows. First, we consider a multivariate generalized linear mixed model for the longitudinal process. Specifically, all components of $\mathbf{Y}_{ik}$ conditional on $\boldsymbol{b}_{ik}$ are independent and the conditional distribution of each $y_{ik}(t_{ijk})$ given $\boldsymbol{b}_{ik}$ is a member of the exponential family with a nonlinear link function given by

$$g_k(E\{y_{ik}(t_{ijk}) | \boldsymbol{b}_{ik}\}) = \eta_{ik}(t_{ijk}, \boldsymbol{b}_{ik}), \quad (2)$$

where $g_k(\cdot)$ is a known monotonic link function and $\eta_{ik}(t, \boldsymbol{b}_{ik})$ is a parametric or nonparametric function of the random effects and t. Also, $\eta_{ik}(t, \boldsymbol{b}_{ik})$ may depend on other covariates of interest, such as gender. Furthermore, we consider a general form of $\eta_{ik}(t, \boldsymbol{b}_{ik})$ given by

$$\eta_{ik}(t, \boldsymbol{b}_{ik}) = \boldsymbol{R}_{ik}(t)\beta_k + \boldsymbol{W}_{ik}(t)\boldsymbol{b}_{ik}, \quad (3)$$

where $\boldsymbol{R}_{ik}(t)$ and $\boldsymbol{W}_{ik}(t)$ are, respectively, the fixed effects and random effects design matrices, and $\boldsymbol{\beta}_k$ and $\boldsymbol{b}_{ik}$ are vectors of the corresponding fixed and random effects parameters.

For the specification in equation (3), we may consider a random varying-coefficient model as follows:

$$\eta_{ik}(t,\boldsymbol{b}_{ik})=\sum_{l=1}^{p} x_{il}\mu_{ikl}(t) \approx \sum_{l=1}^{p}\sum_{s=1}^{S}\theta_{ikl,s}x_{il}\,f_s(t), \quad (4)$$

where $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^T$ is a vector of covariates and $f_s(t)$ are known basis functions, such as B-splines. If $\theta_{ikl,s} = \beta_{kl,s} + b_{ikl,s}$ with $E(b_{ikl,s}) = 0$, then $\boldsymbol{\beta}_k$ includes all coefficients $\beta_{kl,s}$ and $\boldsymbol{b}_{ik}$ includes all $b_{ikl,s}$ for all k, l, and s.

Second, we consider a general multivariate survival model for the survival process as follows. Let $S(t_1, \ldots, t_M|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$ be the joint survivor function of $(T_{i1}, \ldots, T_{iM})$ given $(\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$, where $H_i(t, \boldsymbol{b}_i) = \{\eta_{ik}(\tilde{t}, \mathbf{b}_{ik}) : \tilde{t} \in [0, t), k = 1, \ldots, K\}$ and $\mathbf{z}_i$ is a vector of time-independent covariates. It is assumed that $S(t_1, \ldots, t_M|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$ takes the form

$$F(S_1(t_1|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i), \ldots, S_M(t_M|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i);\boldsymbol{\varphi}), \quad (5)$$

where $F(\ldots)$ is a known function, $\boldsymbol{\varphi}$ is a vector of unknown parameters for characterizing the dependence or association structure, and $S_m(t|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$ for $m = 1, \ldots, M$ are the marginal survival functions given $(\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$. For bivariate time-to-event outcomes, Chi and Ibrahim (2006) have proposed a joint survivor function that is a special case of (5).

For the mth time-to-event outcome, we assume that the marginal hazard function of the ith subject is given by

$$\lambda_m(t|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \mathbf{b}_i)=\lambda_{m0}(t)\exp\{\tilde{g}_{m0}(\mathbf{z}_i;\theta_{m,T})+\sum_{k=1}^{K}\tilde{g}_{mk}(\eta_{ik}(t, \mathbf{b}_{ik}), \mathbf{b}_{ik};\theta_{m,T})\}, \quad (6)$$

where $\lambda_{m0}(t)$ is an unknown baseline hazard function and $\boldsymbol{\theta}_{m,T}$ contains all unknown parameters in $\lambda_m(t|\mathbf{z}_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)$ except $\lambda_{m0}(t)$. Moreover, $\tilde{g}_{mk}(\cdot; \cdot)$ for $k = 0, \ldots, K$ are prespecified functions that characterize the effect of the kth longitudinal profile on the mth time-to-event outcome. Then, we calculate

$S_m(t|z_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)=\exp(-\int_0^t \lambda_m(u|z_i, H_i(t, \boldsymbol{b}_i), \boldsymbol{b}_i)du)$. In the literature, it is common to assume that

$$\tilde{g}_{m0}(\mathbf{z}_i;\theta_{m,T})=\mathbf{z}_i^T\gamma_m \quad \text{and} \quad \tilde{g}_{mk}(\eta_{ik}(t, \boldsymbol{b}_{ik}), \boldsymbol{b}_{ik};\theta_{m,T})=\alpha_{mk}\eta_{ik}(t, \boldsymbol{b}_{ik}) \quad \text{for } k=1, \ldots, K, \quad (7)$$

where $\boldsymbol{\gamma}_m$ is a vector of unknown parameters for the time-independent covariates and α_{mk} are unknown parameters.

Third, we consider a multivariate model for the random effects $\boldsymbol{b}_i$ as follows. Specifically, let $p_0(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$ be a prespecified density function, such as multivariate Gaussian. The density of $\boldsymbol{b}_i$, denoted by $p(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$, is assumed to take the form $p_0(\boldsymbol{b}_i; \boldsymbol{\theta}_b)\psi(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$, where $\psi(\cdot; \cdot)$ is a known and nonnegative function such that $\int p(\boldsymbol{b}_i; \boldsymbol{\theta}_b)d\boldsymbol{b}_i = 1$. For instance, $\psi(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$ can be the square of a polynomial function of individual components of $\boldsymbol{\theta}_b$ and the density of a copula function. We may further consider nonparametric alternatives to the parametric model $p(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$, such as a Dirichlet process.

A formal Bayesian analysis of $(\boldsymbol{\theta}, \boldsymbol{b})$ also involves the specification of a prior distribution $p(\boldsymbol{\theta})$, where $\boldsymbol{b} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$. A typical joint prior specification is to assume $p(\boldsymbol{\theta}) = p(\boldsymbol{\theta}_F)p(\boldsymbol{\theta}_I)$, where $p(\boldsymbol{\theta}_F)$ and $p(\boldsymbol{\theta}_I)$, respectively, denote parametric prior distributions for the

components of $\boldsymbol{\theta}_F$ and non-parametric prior distributions for the components of $\boldsymbol{\theta}_I$. Let $D_O = \{(\mathbf{Y}_i, \mathbf{T}_i, \delta_i): i = 1, \ldots, n\}$. Then, we use Markov chain Monte Carlo (MCMC) methods to obtain samples from the joint posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{b})$, which is given by

$$p(\theta, \boldsymbol{b}|D_o) \propto p(\theta_F)p(\theta_I)\prod_{i=1}^{n}[\,p(\mathbf{Y}_i|\boldsymbol{b}_i;\theta_y)p(\mathbf{T}_i, \delta_i|\boldsymbol{b}_i;\theta_T)p(\boldsymbol{b}_i;\theta_b)]. \quad (8)$$

We focus on nonparametric priors for the nonparametric components in $\eta_{ik}(t, \boldsymbol{b}_{ik})$ and the baseline hazard or cumulative baseline hazard function. We can take different prior distributions, including a Gaussian prior, zero-inflated priors, and stick-breaking priors among others, for the coefficients $\{\theta_{ikl,s}\}$ in model (4). For instance, a Dirichlet process for $\{\theta_{ikl,s}\}$, denoted by $DP(aP_0)$, usually clusters all the longitudinal profiles into one of $k \ll n$ clusters, where P_0 is the base probability measure and a is the confidence parameter. See a nice review of Bayesian nonparametric methods for functional data in Dunson (2009).

Different prior distributions for the baseline hazard $\lambda_{m0}(\cdot)$ or cumulative baseline hazard $\Lambda_{m0}(\cdot)$ include a piecewise constant hazards model, Gamma process model, Beta process model, or a Dirichlet process model. As an illustration, we construct the piecewise constant hazards model for $\lambda_{m0}(\cdot)$. We start with a finite partition of the time axis, $0 < c_{m,1} < \cdots < c_{m,L}$, with $c_{m,L} > T_{im}$ for all i and then set $\lambda_{m0}(t) = h_{m\ell}$ for $t \in I_{m\ell} = (c_{m,\ell-1}, c_{m,\ell}]$. Furthermore, a first-order autoregressive prior or an independent gamma prior can be taken for $\boldsymbol{h}_m = (h_{m1}, \ldots, h_{mL})^T$ (Sinha, 1993; Arjas and Gasbarra, 1994; Ibrahim et al., 2001). An important alternative is the gamma process prior for $\Lambda_{m0}(\cdot)$, that is, $\Lambda_{m0} \sim \mathscr{G}P(c_0\Lambda_m^*, c_0)$, where c_0 is a fixed constant and $\Lambda_m^*(\cdot)$ is a known increasing function with $\Lambda_m^*(0)=0$.

Our aim is to carry out Bayesian inference about parameters of interest, which requires a reasonably "robust" prior $p(\boldsymbol{\theta})$ and the correct specification of $p(\mathbf{Y}_i, \mathbf{T}_i, \delta_i; \boldsymbol{\theta})$. A nonrobust prior $p(\boldsymbol{\theta})$, the presence of outliers, and the misspecification of the JMLS may introduce serious bias in the estimation and inference on $\boldsymbol{\theta}$. Thus, it is crucial to assess the sensitivity of statistical inference to the prior, the sampling distribution, and outliers. We note that existing frequentist diagnostic tools are not sufficient for this endeavor (Dobson and Henderson, 2003; Rizopoulos et al., 2008; Rizopoulos and Ghosh, 2011).

Example 1

For the purposes of illustration, we consider an example with two longitudinal markers and bivariate survival times. In this case, $K = M = 2$. Specifically, each longitudinal response is given by

$$\begin{aligned}
y_{ik}(t_{ijk}) &= \eta_{ik}(t_{ijk}, \boldsymbol{b}_{ik}) + \varepsilon_{ijk} \\
&= \beta_{k0} + \beta_{k1}t_{ijk} + \beta_{k2}r_i + b_{ik0} + b_{ik1}t_{ijk} + \varepsilon_{ijk},
\end{aligned} \quad (9)$$

for $i = 1, \ldots, 100$, $k = 1, 2$, and $j = 1, \ldots, n_i$, where the r_is represent a baseline covariate in the longitudinal model. Moreover, it is assumed that $t_{ij1} = t_{ij2}$ for all i and j, $\boldsymbol{e}_{ij} = (\varepsilon_{ij1}, \varepsilon_{ij2})^T$ are independently and identically distributed as $N_2(\mathbf{0}, \boldsymbol{\Sigma})$, and the random effects $\boldsymbol{b}_i=(\boldsymbol{b}_{i1}^T, \boldsymbol{b}_{i2}^T)^T$ are distributed as $N_4(\mathbf{0}, \boldsymbol{\Phi})$, where $\boldsymbol{b}_{ik} = (b_{ik0}, b_{ik1})^T$ for $k = 1, 2$. Here $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}$ for $k = 1, 2$ are covariance matrices. Conditional on $\boldsymbol{b}_i$, the two event and censoring times are assumed to be independent and their marginal hazard functions are given by

$$\lambda_m(t|\boldsymbol{b}_i, \boldsymbol{z}_i) = \lambda_{m0}(t)\exp\left\{\alpha_{m1}\eta_{i1}(t, \boldsymbol{b}_{i1}) + \alpha_{m2}\eta_{i2}(t, \boldsymbol{b}_{i2}) + \boldsymbol{z}_i^T \boldsymbol{\gamma}_m\right\} \quad (10)$$

for $m = 1, 2$, where $\boldsymbol{z}_i = (z_{i1}, z_{i2})^T$ is a vector of time-independent covariates. Let $\boldsymbol{Y}_i(t) = (y_{i1}(t), y_{i2}(t))^T$ and $\boldsymbol{\eta}_i(t, \boldsymbol{b}_i) = (\eta_{i1}(t, \boldsymbol{b}_{i1}), \eta_{i2}(t, \boldsymbol{b}_{i2}))^T$. The density of $(\boldsymbol{Y}_i, \mathbf{T}_i, \delta_i, \boldsymbol{b}_i)$ given $\boldsymbol{\theta}$ for the ith subject, denoted by $p(\boldsymbol{Y}_i, \mathbf{T}_i, \delta_i, \boldsymbol{b}_i; \boldsymbol{\theta})$, is given by

$$C_0 |\boldsymbol{\Phi}|^{-1/2} \exp\left(-\tfrac{1}{2}\boldsymbol{b}_i^T \boldsymbol{\Phi}^{-1} \boldsymbol{b}_i\right) \prod_{m=1}^{2} \lambda_m \left(T_{im}|\boldsymbol{b}_i, \boldsymbol{z}_i\right)^{\delta_{im}}$$
$$\times \exp\left\{-\int_0^{T_{im}} \lambda_m(u|\boldsymbol{b}_i, \boldsymbol{z}_i)du\right\}$$
$$\times \prod_{j=1}^{n_i} \left(|\textstyle\sum|^{-1/2} \exp\left[-\tfrac{1}{2}\{\boldsymbol{Y}_i(t_{ij}) - \eta_i(t_{ij}, b_i)\}^T \times \textstyle\sum^{-1}\{\boldsymbol{Y}_i(t_{ij}) - \eta_i(t_{ij}, b_i)\}\right]\right),$$

where C_0 is a constant independent of $\boldsymbol{\theta}$.

To carry out a Bayesian analysis, we take a joint prior distribution for $\boldsymbol{\theta}$ as follows:

$$\alpha_m = (\alpha_{m1}, \alpha_{m2})^T \sim N\left(\alpha_m^0, \boldsymbol{H}_\alpha^0\right), \gamma_m \sim N\left(\gamma_m^0, \boldsymbol{H}_\gamma^0\right),$$
$$\textstyle\sum^{-1} \sim \text{Wishart}_2\left(\boldsymbol{R}^0, \rho^0\right), \boldsymbol{\Phi}^{-1} \sim \text{Wishart}_4\left(\boldsymbol{R}_\varphi^0, \rho^0\right) \quad (11)$$
$$\beta_k = (\beta_{k0}, \beta_{k1}, \beta_{k2})^T \sim N\left(\beta_k^0, \boldsymbol{H}_\beta^0\right),$$

for $k, m = 1, 2$, where $\alpha_m^0, \boldsymbol{H}_\alpha^0, \gamma_m^0, \boldsymbol{H}_\gamma^0, \boldsymbol{R}^0, \rho^0, \boldsymbol{R}_\varphi^0$, and ρ^0 are prespecified hyperparameters. For the baseline hazard $\lambda_{m0}(\cdot)$, we take a piecewise constant hazards model with 250 subintervals with equal lengths such that $\lambda_{m0}(t) = \sum_{l=1}^{L} h_{ml}\mathbf{1}(t \in (c_{m,l-1}, c_{m,l}])$, where the $c_{m, l}$s are prespecified constants. We take $h_{m l} \sim \Gamma(\tau_{0l}, \tau_{1l})$ for $l = 1, \dots, 250$ and $m = 1, 2$. We use MCMC methods to conduct our Bayesian influence analysis on $\boldsymbol{\theta}$ and $\boldsymbol{b}$.

3. Bayesian Influence Analysis

We address three issues related to Bayesian influence analysis of JMLS: perturbation models for perturbing the JMLS, appropriate perturbations, and Bayesian influence measures.

3.1. Perturbation Models and Appropriate Perturbations

We introduce three classes of perturbation models to formally perturb JMLS. Let $\boldsymbol{\omega}$ be a perturbation vector in a set $\Omega \subset R^W$, which represents a Euclidean space of dimension W, where W is an integer. The perturbed model $\mathcal{M} = \{p(D_o, \boldsymbol{b}; \boldsymbol{\theta}, \boldsymbol{\omega}): \boldsymbol{\omega} \in \Omega\}$ characterizes various perturbations to the assumed density of $p(D_o, \boldsymbol{b}; \boldsymbol{\theta})$ such that $\int p(D_o, \boldsymbol{b}; \boldsymbol{\theta}, \boldsymbol{\omega})dD_o\, d\boldsymbol{b} = 1$ and $p(D_o, \boldsymbol{b}; \boldsymbol{\theta}, \boldsymbol{\omega}^0) = p(D_o, \boldsymbol{b}; \boldsymbol{\theta})$ for a unique $\boldsymbol{\omega}^0$, which represents no perturbation. The first class of perturbations is to individually perturb a subject's longitudinal profile, repeated measures within each subject, covariates, survival time, and censoring indicator. For instance, we introduce a perturbation vector $\boldsymbol{\omega}_{y,i}$ into $p(\boldsymbol{Y}_i|\boldsymbol{b}_i; \boldsymbol{\theta}_y)$ to perturb $\boldsymbol{Y}_i$ such that the perturbed density is given by

$$\prod_{i=1}^{n} \{p(\mathbf{T}_i, \delta_i|\boldsymbol{b}_i; \theta_T)p(\boldsymbol{b}_i; \theta_b)p(\boldsymbol{Y}_i|\boldsymbol{b}_i; \theta_y, \omega_{y,i})\}, \quad (12)$$

in which $\boldsymbol{\omega} = \{ \boldsymbol{\omega}_{y,1}, \ldots, \boldsymbol{\omega}_{y,n} \}$. This single-case perturbation equation (12) to the individual's longitudinal profile is primarily designed to detect one or a few influential subjects, whose longitudinal profiles differ significantly from the other subjects in the evolution process. Furthermore, we may introduce a perturbation vector $\boldsymbol{\omega}_i$ into $p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \boldsymbol{\theta}_T)$ to perturb $(\mathbf{T}_i, \delta_i)$ such that the perturbed density is given by

$$\prod_{i=1}^{n} \{ p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \theta_T, \omega_{T,i}) p(\boldsymbol{b}_i; \theta_b) p(\boldsymbol{Y}_i | \boldsymbol{b}_i; \theta_y) \}, \quad (13)$$

in which $\omega = (\omega_{T,1}^T, \ldots, \omega_{T,n}^T)^T$. This single-case perturbation equation (13) to the survival data is designed to reveal influential values of survival times relative to other survival times.

The second class is to perturb the shared random effects that underly both the longitudinal measurement and survival processes. For instance, we introduce a perturbation vector $\boldsymbol{\omega}_{b,i}$ to simultaneously perturb $(\mathbf{Y}_i, \mathbf{T}_i, \boldsymbol{\delta}_i)$ with the presence of $\boldsymbol{b}_i$ such that

$$\prod_{i=1}^{n} \{ p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \theta_T, \omega_{b,i}) p(\boldsymbol{b}_i; \theta_b) p(\boldsymbol{Y}_i | \boldsymbol{b}_i; \theta_y, \omega_{b,i}) \}. \quad (14)$$

In equation (14), we use the same $\boldsymbol{\omega}_{b,i}$ to simultaneously perturb $p(\mathbf{Y}_i | \boldsymbol{b}_i; \boldsymbol{\theta}_y)$ and $p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \boldsymbol{\theta}_T)$ to delineate large discrepancies between the longitudinal profile and the corresponding survival time, each of which may not be influential by themselves. This single-case perturbation equation (14) is primarily designed to detect influential survival times, whose occurrence is low relative to their subject-specific longitudinal profiles. An alternative perturbation is to introduce a perturbation $\boldsymbol{\omega}_{b,i}$ into $p(\boldsymbol{b}_i; \boldsymbol{\theta}_b)$ such that

$$\prod_{i=1}^{n} \{ p(\mathbf{T}_i, \delta_i | \boldsymbol{b}_i; \theta_T) p(\boldsymbol{b}_i; \theta_b, \omega_{b,i}) p(\boldsymbol{Y}_i | \boldsymbol{b}_i; \theta_y) \}. \quad (15)$$

Perturbation equation (15) is designed to detect "influential" random effects $\boldsymbol{b}_i$, whereas their corresponding $\mathbf{T}_i$ and $\mathbf{Y}_i$ may be not influential. Furthermore, by setting $\boldsymbol{\omega}_{b,i} = \boldsymbol{\omega}_b$, we can formally assess the parametric distributional assumptions of the random effects $\boldsymbol{b}_i$ and the amount of such perturbations to statistical inferences, such as their impact on parameter estimation.

The third class includes perturbations to the prior $p(\boldsymbol{\theta})$ and the simultaneous perturbations to all three components of the Bayesian model. A fundamental issue associated with any Bayesian analysis is how much posterior quantities, such as the Bayes factor, parameter estimates, and credible (or highest posterior density) intervals, are sensitive to changes in the prior distribution. Thus, it is important to assess both Bayesian semiparametric assumptions for the longitudinal profiles and perturbations regarding the nonparametric prior processes for the cumulative baseline hazard function. For instance, we consider a prior perturbation to $\Lambda_{m0}(t) \sim \mathcal{G}P(c_0 \Lambda^*(t), c_0)$ by assuming $\Lambda_{m0}(t) \sim \mathcal{G}P(c_0 \Lambda^*(t, \boldsymbol{\omega}_P), c_0)$ such that $\Lambda^*(t, \omega_P) = \int_0^t k_*(s, \omega_P) h_0(s) ds$ and $\Lambda^*(t, \mathbf{0}) = \Lambda^*(t)$. Combining the first two classes with the third class, we can obtain various simultaneous perturbations to the prior, the sampling distribution, and the data, which allows us to assess the simultaneous sensitivity of all components of a Bayesian analysis.

Example 1 (continued)—We consider some simultaneous perturbations as follows:

$$
\begin{aligned}
Y_i(t_{ij}) &= \eta_i(t_{ij}, \boldsymbol{b}_i) + \omega_{y,i}\mathbf{1}_2 + \boldsymbol{\varepsilon}_{ij}, \\
p(\boldsymbol{b}_i; \theta_b, \omega_{b,i}) &\sim N_4(\mathbf{0}, \omega_{b,i}^{-1}\boldsymbol{\Phi}), \\
p(\alpha_m | \omega_{\alpha 1}, \omega_{\alpha 2}) &\sim N(\alpha_m^0 + \omega_{\alpha 2}\mathbf{1}_2, \boldsymbol{H}_\alpha^0/\omega_{\alpha 1}), \\
p(\gamma_m | \omega_{\gamma 1}, \omega_{\gamma 2}) &\sim N(\gamma_m^0 + \omega_{\gamma 2}\mathbf{1}_2, \boldsymbol{H}_\gamma^0/\omega_{\gamma 1}).
\end{aligned}
\quad (16)
$$

Thus, the perturbed (unnormalized) log-posterior is given by

$$
\begin{aligned}
&-0.5 \sum_{i=1}^{n} \sum_{j=1}^{n_i} \{Y_i(t_{ij}) - \eta_i(t_{ij}, \boldsymbol{b}_i) - \omega_{y,i}\mathbf{1}_2\}^T \\
&\times \Sigma^{-1} \{Y_i(t_{ij}) - \eta_i(t_{i,j}, \boldsymbol{b}_i) - \omega_{y,i}\mathbf{1}_2\} \\
&-0.5 \left\{ \sum_{i=1}^{n} \omega_{b,i} \boldsymbol{b}_i^T \boldsymbol{\Phi}^{-1} \boldsymbol{b}_i^T + \sum_{i=1}^{n} \log|\omega_{b,i}^{-1}\boldsymbol{\Phi}| \right. \\
&+ (\beta_k - \beta_k^0)^T (\boldsymbol{H}_\beta^0)^{-1} (\beta_k - \beta_k^0) + \log|\boldsymbol{H}_\beta^0|\} \\
&-0.5 \left\{ \sum_{m=1}^{2} \omega_{\alpha 1} (\alpha_m - \alpha_m^0 - \omega_{\alpha 2}\mathbf{1}_2)^T \right. \\
&\times \boldsymbol{H}_\alpha^{0-1} (\alpha_m - \alpha_m^0 - \omega_{\alpha 2}\mathbf{1}_2) + 2\log|\omega_{\alpha 1}^{-1}\boldsymbol{H}_\alpha^0| \\
&+ 2\log|\omega_{\gamma 1}^{-1}\boldsymbol{H}_\gamma^0| + \sum_{m=1}^{2} \omega_{\gamma 1} (\gamma_m - \gamma_m^0 - \omega_{\gamma 2}\mathbf{1}_2)^T \\
&\times \boldsymbol{H}_\gamma^{0-1} (\gamma_m - \gamma_m^0 - \omega_{\gamma 2}\mathbf{1}_2) \}.
\end{aligned}
\quad (17)
$$

In this case, $\boldsymbol{\omega}$ contains $(\omega_{y,1}, \ldots, \omega_{y,n}, \omega_{b,1}, \ldots, \omega_{b,n}, \omega_{\alpha 1}, \omega_{\alpha 2}, \omega_{\gamma 1}, \omega_{\gamma 2})^T$ and $\boldsymbol{\omega}^0 = (0, \ldots, 0, 1, \ldots, 1, 1, 0, 1, 0)^T$ represents no perturbation.

After perturbing the JMLS, we need to quantify the amount of perturbation introduced by each perturbation, the extent to which each component of a perturbation vector contributes to, and the degree of orthogonality for the components of a perturbation vector (Amari, 1990; Zhu, Ibrahim, and Tang, 2011). This is very critical for us to properly pinpoint the cause (e.g., prior) of a large effect. Specifically, we regard the perturbed model $p(D_o, \boldsymbol{b}, \boldsymbol{\theta}; \boldsymbol{\omega})$ as the probability density of $(D_o, \boldsymbol{b}, \boldsymbol{\theta})$ for $\boldsymbol{\omega}$ and then calculate its score function $\partial_{\omega_h}\ell(\boldsymbol{\omega})$, where $\partial_{\omega_h} = \partial/\partial\omega_h$, $\ell(\boldsymbol{\omega}) = \log p(D_o, \boldsymbol{b}, \boldsymbol{\theta}; \boldsymbol{\omega})$, and ω_h is the hth component of $\boldsymbol{\omega}$. Thus, the Fisher information matrix with respect to $\boldsymbol{\omega}$, denoted by $\boldsymbol{G}(\boldsymbol{\omega}) = (g_{hs}(\boldsymbol{\omega}))$, is a $W \times W$ matrix with its (h, s) element given by $g_{hs}(\omega) = E_\omega[\partial_{\omega_h}\ell(\omega)\partial_{\omega_s}\ell(\omega)] = -E_\omega[\partial_{\omega_h\omega_s}^2\ell(\omega)]$, for $h, s = 1, \ldots, W$, where E_ω denotes the expectation taken with respect to $p(D_o, \boldsymbol{b}, \boldsymbol{\theta}; \boldsymbol{\omega})$.

We call $p(D_o, \boldsymbol{b}, \boldsymbol{\theta}; \boldsymbol{\omega})$ an *appropriate perturbation* model if $\boldsymbol{G}(\boldsymbol{\omega}^0)$ equals $c\mathbf{I}_W$, where c is any positive scalar and $\mathbf{I}_W$ is a $W \times W$ identity matrix (Zhu et al., 2007). Specifically, $g_{hh}(\boldsymbol{\omega})$ can be regarded as the amount of perturbation introduced by ω_h, whereas the correlation $g_{hs}(\omega)/\sqrt{g_{hh}(\omega)g_{ss}(\omega)}$ indicates an association between ω_h and ω_s. The diagonal structure of $\boldsymbol{G}(\boldsymbol{\omega})$ implies that all components of $\boldsymbol{\omega}$ are orthogonal to each other. Orthogonal subcomponents of $\boldsymbol{\omega}$ allow for easy detection of the cause of a large effect. If $\boldsymbol{G}(\boldsymbol{\omega}^0)$ is not diagonal, then we choose a new perturbation vector $\tilde{\boldsymbol{\omega}}$, defined by

$$
\tilde{\omega}(\omega) = \omega^0 + \boldsymbol{G}(\omega^0)^{1/2}(\omega - \omega^0). \quad (18)
$$

Based on $\tilde{\boldsymbol{\omega}}$, we can obtain a new perturbation model $p(D_o, \boldsymbol{b}, \boldsymbol{\theta}; \tilde{\boldsymbol{\omega}})$ such that $\boldsymbol{G}(\tilde{\boldsymbol{\omega}})$ evaluated at $\boldsymbol{\omega}^0$ equals $\boldsymbol{I}_W$.

Example 1 (continued)—Let $g_{y,ii}=n_i E(\mathbf{1}_2^T \sum^{-1} \mathbf{1}_2)=n_i \rho^0 \mathbf{1}_2^T R^0 \mathbf{1}_2$ for $i = 1, \ldots, n$. It follows from the definition of $G(\boldsymbol{\omega}^0)$ that

$G(\omega^0)=\mathrm{diag}(g_{y,11}, \ldots, g_{y,nn}, 2\mathbf{1}_n, M, \mathbf{1}_2^T H_\alpha^{0-1} \mathbf{1}_2, M, \mathbf{1}_2^T H_\gamma^{0-1} \mathbf{1}_2)$. The diagonal structure of $G(\boldsymbol{\omega}^0)$ indicates that all components of $\boldsymbol{\omega}$ are orthogonal to each other. However, because $G(\boldsymbol{\omega}^0)$ does not take the form of $c\mathbf{I}_W$, this indicates that different components of $\boldsymbol{\omega}$ introduce different amounts of perturbations. For instance, a large $g_{y,ii}$ indicates that $\omega_{y,i}$ introduces a large perturbation for a subject with more repeated measures (n_i). In practice, we can always choose the appropriate perturbation scheme in equation (18).

3.2. Local Influence Measures

Let $f(p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega})) = f(\boldsymbol{\omega})$ be a real objective function (e.g., Bayes factor) of the perturbed model, which defines the aspect of inference of interest for sensitivity analysis. We use $f(\boldsymbol{\omega})$ to measure the effect of a small perturbation $\boldsymbol{\omega}$ to a JMLS around $\boldsymbol{\omega}^0$. Specifically, we consider a smooth curve of the perturbed model $p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega}(t))$ such that $p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega}(0)) = p(D_o, \mathbf{b}, \boldsymbol{\theta})$. Then, the score function of $\ell(\boldsymbol{\omega}(t))$ with respect to t is equal to $\partial_t \ell(\boldsymbol{\omega}(t)) = \mathbf{h}^T \partial_\omega \ell(\boldsymbol{\omega}(t))$, where $\partial_t = \partial/\partial t$, $\partial_\omega = \partial/\partial \omega$, and $\partial_t \boldsymbol{\omega}(t)|_{t=0} = \mathbf{h} \in R^W$. Then, we quantify the effects of introducing $\boldsymbol{\omega}(t)$ to perturb the JMLS by using $\{f(\boldsymbol{\omega}(t)) - f(\boldsymbol{\omega}(0))\}^2$ relative to the Kullback–Leibler divergence between $p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega}(0))$ and $p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega}(t))$, denoted by $S(\boldsymbol{\omega}(0), \boldsymbol{\omega}(t))$.

For small t, it follows from a Taylor's series expansion that $S(\boldsymbol{\omega}(0), \boldsymbol{\omega}(t)) \approx 0.5 t^2 \mathbf{h}^T G(\boldsymbol{\omega}^0)\mathbf{h}$ and $f(\boldsymbol{\omega}(t)) - f(\boldsymbol{\omega}(0)) = \dot{f}_\mathbf{h}(0)t + O(t^2)$, where $\dot{f}_\mathbf{h}(0)=\nabla_f^T \mathbf{h}$, in which $\nabla_f = \partial_\omega f(\boldsymbol{\omega}^0)$. If $\nabla_f \neq 0$, we use a quantity $\mathrm{FI}_{f,\mathbf{h}}$ called the *first-order influence measure* (FI) in the direction $\boldsymbol{h} \in R^W$ for the objective function $f(\boldsymbol{\omega})$, which is given by

$$
\begin{aligned}
\mathrm{FI}_{f,\mathbf{h}}=\mathrm{FI}_{f(\omega^0),\mathbf{h}} &=\lim_{t\to 0}\frac{\{f(\omega(t))-f(\omega(0))\}^2}{2S(\omega(0),\omega(t))}\\
&=\mathrm{FI}_{f(\omega^0),\mathbf{h}}=\frac{\mathbf{h}^T \nabla_f \nabla_f^T \mathbf{h}}{\mathbf{h}^T G\mathbf{h}},
\end{aligned} \tag{19}
$$

where $G = G(\boldsymbol{\omega}^0)$. For the appropriate perturbation $\tilde{\boldsymbol{\omega}}(\boldsymbol{\omega})$ given in equation (18), $\mathrm{FI}_{f,\mathbf{h}}$ reduces to $\mathbf{h}^T G^{-1/2} \nabla_f \nabla_f^T G^{-1/2}\mathbf{h}$ with the constraint $\mathbf{h}^T \mathbf{h} = 1$.

We use the maximum value of $\mathrm{FI}_{f,\mathbf{h}}$ and its associated eigenvector, denoted by $\mathbf{h}_{\max}$, as influence measures to quantify the largest degree and influential direction of local influence of $\tilde{\boldsymbol{\omega}}$ to the JMLS. It can be easily shown that $\mathrm{FI}_{f,\mathbf{h}_{\max}}$ equals $\nabla_f^T G^{-1} \nabla_f$ and $\mathbf{h}_{\max}=G^{-1/2}\nabla_f / \sqrt{\nabla_f^T G^{-1}\nabla_f}$. In particular, $\boldsymbol{h}_{\max}$ can be used to detect robustness of priors or identify influential observations and incorrect sampling distributional assumptions for single-case and global perturbations (Cook, 1986). Following Zhu and Lee (2001) and Zhu et al. (2007), we also suggest inspecting $\mathrm{FI}_{f,\,\mathbf{e}_i}$ to identify the most significant components of $\tilde{\boldsymbol{\omega}}$, where $\mathbf{e}_i$ is a $W \times 1$ vector with a 1 for the ith element and 0 otherwise. For instance, we consider the Bayes factor given by $f(\boldsymbol{\omega}) = \log p(D_o; \boldsymbol{\omega}) - \log p(D_o; \boldsymbol{\omega}^0)$, where $p(D_o; \boldsymbol{\omega}) = \int p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega})d\mathbf{b}d\boldsymbol{\theta}$. Thus, under some smoothness conditions, it can be shown that $\nabla_f = E_{\omega^0}[\partial_\omega \log p(D_o, \mathbf{b}, \boldsymbol{\theta}; \boldsymbol{\omega}^0)|D_o]$ and $\boldsymbol{h}_{\max}=G^{-1/2}\nabla_f / \sqrt{\nabla_f^T G^{-1}\nabla_f}$. To calculate the local influence measures associated with $f(\boldsymbol{\omega})$, we just need to compute ∇_f and G. In practice, we can use MCMC methods to draw samples $\{(\boldsymbol{\theta}^{(s)}, \mathbf{b}^{(s)}): s = 1, \ldots, S_0\}$ from $p(\boldsymbol{\theta}, \mathbf{b}; D_o, \boldsymbol{\omega}^0)$ to approximate ∇_f via $\nabla_f \approx S_0^{-1} \sum_{s=1}^{S_0} \partial_\omega \log p(D_o, \mathbf{b}^{(s)}, \theta^{(s)}; \omega^0)$.

We can also carry out Bayesian local influence when $\nabla_f = 0$. Because $f(\boldsymbol{\omega}(t)) = f(\boldsymbol{\omega}(0)) + 0.5\mathbf{h}^T \boldsymbol{H}_f \mathbf{h} t^2 + O(t^3)$, where $\boldsymbol{H}_f = \partial_\omega^2 f(\omega^0)$, we introduce a *second-order influence measure* (SI) in the direction $\mathbf{h} \in R^W$, given by

$$\text{SI}_{f,\mathbf{h}} = \text{SI}_{f(\omega^0),\mathbf{h}} = \lim_{t \to 0} \frac{\{f(\omega(t)) - f(\omega(0))\}}{S(\omega(0), \omega(t))} = \frac{\mathbf{h}^T \boldsymbol{H}_f \mathbf{h}}{\mathbf{h}^T \boldsymbol{G} h}. \quad (20)$$

For $\tilde{\boldsymbol{\omega}}(\boldsymbol{\omega})$ in equation (18), $\text{SI}_{f,\mathbf{h}}$ reduces to $\mathbf{h}^T \boldsymbol{G}^{-1/2} \boldsymbol{H}_f \boldsymbol{G}^{-1/2} \mathbf{h}$, where $\mathbf{h}^T \mathbf{h} = 1$. Moreover, we also use SI_{f,e_i} and the eigenvalue–eigenvector pairs of $\boldsymbol{G}^{-1/2} \boldsymbol{H}_f \boldsymbol{G}^{-1/2}$ as our influence measures. We use the eigenvector of the largest eigenvalue of $\boldsymbol{G}^{-1/2} \boldsymbol{H}_f \boldsymbol{G}^{-1/2}$, denoted by $\mathbf{h}_{\max}$, as influence measures to quantify the influential direction of local influence of $\tilde{\boldsymbol{\omega}}$ to the JMLS.

Finally, we examine the influence measures associated with three common objective functions, these being the φ–divergence, the posterior mean distance, and the Bayes factor, and include the detailed formulas in the supplementary document. Although all three objective functions can assess the local influence of a perturbation vector $\boldsymbol{\omega}$ to the JMLS, there is a conceptual difference among these measures. The φ–divergence and the Bayes factor quantify the effects of introducing $\boldsymbol{\omega}$ on the overall posterior distribution, whereas the posterior mean distance quantifies the effect of $\boldsymbol{\omega}$ on the posterior mean of $\boldsymbol{\theta}$. Because the perturbation vector $\boldsymbol{\omega}$ may influence various characteristics of the posterior distribution, such as the shape, mode, and mean, the φ–divergence, and the Bayes factor can be more sensitive to some perturbations of the posterior distribution compared to the posterior mean for certain perturbation schemes $\boldsymbol{\omega}$. In contrast, the posterior mean distance may be more sensitive to perturbations which have a dramatic effect on the posterior mean.

4. Application to the IBCSG Data

We applied our proposed methodology to both simulated data and the IBCSG data discussed in the Introduction Section. For the sake of space, we only present some influence analysis results for the IBCSG data here. We refer the reader to the supplementary document for further details.

For the IBCSG data, we considered a JMLS for jointly investigating the relationship between the multidimensional QOL and the bivariate failure time variables DFS and OS. To satisfy the normality assumption for the four considered longitudinal QOL indicators (appetite, denoted as y_1; perceived coping, denoted as y_2; mood, denoted as y_3; and physical well-being, denoted as y_4), we transformed their corresponding observed values of QOL to $\sqrt{100 - \text{QOL}}$ (Chi and Ibrahim, 2006). The transformed QOLs decreased over time and were scaled between 0 and 10 with smaller values reflecting better QOL. There were 832 patients from Switzerland, Sweden, and New Zealand/Australia with a total of 2154 QOL observations being included in this analysis.

Let $y_{i1}(t_{ij1}), \ldots, y_{i4}(t_{ij4})$ be the observed values of the transformed QOLs for the ith patient at the jth time point, respectively. We considered the following JMLS:

$$\lambda_m(t|\boldsymbol{b}_i, \boldsymbol{z}_i) = \lambda_{m0}(t)\exp\{\alpha_{m1}\eta_{i1}(t, \boldsymbol{b}_{i1}) + \cdots + \alpha_{m4}\eta_{i4}(t, \boldsymbol{b}_{i4}) + \boldsymbol{z}_i^T \gamma_m\}$$
$$y_{ik}(t_{ijk}) = \eta_{ik}(t_{ijk}, \boldsymbol{b}_{ik}) + \varepsilon_{ijk} = \beta_{k0} + \beta_{k1}x_{i1} + \cdots + \beta_{k6}x_{i6} + \beta_{k7}t_{ijk} + b_{ik0} + b_{ik1}t_{ijk} + \varepsilon_{ijk}, \quad (21)$$

for $i = 1, \ldots, 832$, $j = 1, 2, 3$, $k = 1, \ldots, 4$, and $m = 1, 2$, where $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{i6})^T$ includes the number of initial cycles, reintroduction, interaction of the number of initial cycles and reintroduction, age, residency for Switzerland, and residency for Sweden. Moreover, $\boldsymbol{z}_i$

includes the number of initial cycles, reintroduction, interaction of the number of initial cycles and reintroduction, age, number of positive nodes, and ER status. We assumed that $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij1}, \ldots, \varepsilon_{ij4})^T$ are independently and identically distributed as $N_4(\boldsymbol{0}, \boldsymbol{\Sigma})$, and the random effects $\boldsymbol{b}_{ik} = (b_{ik0}, b_{ik1})^T$ are independently and identically distributed as $N_2(\boldsymbol{0}, \boldsymbol{\varphi}_k)$ for $i = 1, \ldots, 832$, $j = 1, 2, 3$, and $k = 1, \ldots, 4$. For the baseline hazard $\lambda_{m0}(\cdot)$, we take the piecewise constant hazards model with 250 subintervals with equal lengths such that

$$\lambda_{m0}(t) = \sum_{l=1}^{L} h_{ml} \mathbf{1}(t \in (c_{m,l-1}, c_{m,l}]) \text{ with } L = 250, \text{ where the } c_{m,l}\text{s are prespecified constants.}$$

To conduct a Bayesian analysis, we specified the following prior distributions:

$$\begin{aligned} \alpha_m = (\alpha_{m1}, \ldots, \alpha_{m4})^T &\sim N\left(\alpha_m^0, \boldsymbol{H}_\alpha^0\right), \gamma_m \sim N\left(\gamma_m^0, \boldsymbol{H}_\gamma^0\right), \Sigma^{-1} \sim \text{Wishart}_4(\boldsymbol{R}^0, \rho^0), \\ \beta_k = (\beta_{k0}, \ldots, \beta_{k7})^T &\sim N\left(\beta_k^0, \boldsymbol{H}_\beta^0\right), h_{ml} \sim \Gamma(\tau_{0l}, \tau_{1l}), \boldsymbol{\Phi}_k^{-1} \sim \text{Wishart}_2\left(\boldsymbol{R}_{\varphi k}^0, \rho_{\varphi k}^0\right) \end{aligned} \quad (22)$$

for $m = 1$ and 2, and $k = 1, \ldots, 4$, where $\boldsymbol{h} = \{h_{ml}: m = 1, 2, l = 1, \ldots, L\}$, $\alpha_m^0, \boldsymbol{H}_\alpha^0, \gamma_m^0, \boldsymbol{H}_\gamma^0$, $\boldsymbol{R}^0, \rho^0, \beta_k^0, \boldsymbol{H}_\beta^0, \tau_{0l}, \tau_{1l}, \boldsymbol{R}_{\varphi k}^0$, and $\rho_{\varphi k}^0$ are prespecified hyperparameters. Moreover, $\alpha_m^0, \gamma_m^0, \beta_k^0, \boldsymbol{R}^0$, and $\boldsymbol{R}_{\varphi k}^0$ were set to their Bayesian posterior means obtained from MCMC methods based on noninformative prior distributions for $\boldsymbol{a}_m, \boldsymbol{\gamma}_m, \Sigma^{-1}, \boldsymbol{\beta}$, and $\boldsymbol{\varphi}_k$. We used MCMC methods, whose key steps are described in the supplementary document, to carry out the Bayesian analysis.

To illustrate our influence analysis, we considered five different perturbations to the JMLS and carried out the associated influence analysis. Specifically, for each perturbation $\boldsymbol{\omega}$, we calculated the metric tensor $\boldsymbol{G}$ and then took the new appropriate perturbation $\tilde{\boldsymbol{\omega}}(\boldsymbol{\omega})$ given in equation (18). Detailed derivations of influence quantities, such as $\boldsymbol{G}(\boldsymbol{\omega}^0)$, can be found in the supplementary document. We chose the Bayes factor as the objective function and then calculated the local influence measure $\boldsymbol{h}_{\max}$ of the Bayes factor for each perturbation scheme by using the MCMC output. Specifically, a total of 5,000 iterations after 5,000 burn-in samples were used to compute all local influence measures.

The first perturbation is a single-case perturbation, which is obtained by perturbing each subject's longitudinal profile as follows:

$$y_{ik}(t_{ijk}, \omega_{ijk}) = \beta_{k0} + \beta_{k1} x_{i1} + \cdots + \beta_{k6} x_{i6} + \beta_{k7} t_{ijk} + b_{ik0} + b_{ik1} t_{ijk} + \varepsilon_{ijk}/\omega_{ijk}.$$

In this case, $\omega = (\omega_{11}^T, \ldots, \omega_{1n_1}^T, \ldots, \omega_{n1}^T, \ldots, \omega_{nn_n}^T)^T$, in which $\boldsymbol{\omega}_{ij} = (\omega_{ij1}, \ldots, \omega_{ij4})^T$ for $i = 1, \ldots, n = 832$, and $j = 1, \ldots, 3$, and $\boldsymbol{\omega}^0 = \boldsymbol{0}$ presents no perturbation. This single-case perturbation is designed to detect influential transformed QOLs of the longitudinal profiles in the evolution process. Figure 1 presents the subcomponents of $\boldsymbol{h}_{\max}$ corresponding to all cases (i, j) for $k = 1, 2, 3$, and 4, where (i, j) represents the ith patient at the jth time point. Inspecting Figure 1 reveals the most influential cases as (158,2), (306,1), (619,3), (788,2), (802,2), and (803,2) for $k = 1$; (39,2), (541,1), and (625,2) for $k = 2$; (103,2), (130,2), (209,1), (331,2), and (331,3) for $k = 3$; and (53,2), (103,2), (208,2), (306,1), (530,1), (780,2), and (792,3) for $k = 4$.

The second perturbation is also a single-case perturbation, which perturbs each marginal hazard function as follows:

$$\lambda_m(t|\boldsymbol{b}_i, \boldsymbol{z}_i, \omega_{im}) = \lambda_{m0}(t) \exp\{\alpha_{m1}\eta_{i1}(t, \boldsymbol{b}_i) + \cdots + \alpha_{m4}\eta_{i4}(t, \boldsymbol{b}_i) + z_i^T \gamma_m + \omega_{im}\}$$

for $i = 1, \ldots, n$ and $m = 1, 2$. In this case, $\boldsymbol{\omega} = (\omega_{11}, \omega_{21}, \ldots, \omega_{n1}, \omega_{n2})$ with $\boldsymbol{\omega}^0 = \mathbf{1}$ representing no perturbation. This single-case perturbation is used to detect influential disease-free survival and overall survival times in the survival process. Inspecting Figure 2 reveals patients 15, 18, and 28 as most influential for $m = 1$, and patients 15, 42, 85, 96, 136, 234, 281, and 282 for $m = 2$ by our local influence measure $\boldsymbol{h}_{\max}$ for the Bayes factor.

The third perturbation is to simultaneously perturb the shared random effects $\mathbf{b}_i$ in both the longitudinal profiles and the marginal hazard functions:

$$y_{ik}(t_{ijk}, \omega_{ik}) = \beta_{k0} + \beta_{k1}x_{i1} + \cdots + \beta_{k6}x_{i6} + \beta_{k7}t_{ijk} + \omega_{ik}(b_{ik0} + b_{ik1}t_{ijk}) + \varepsilon_{ijk}$$
$$\triangleq \eta_{ik}(t_{ijk}, \boldsymbol{b}_{ik}, \omega_{ik}) + \varepsilon_{ijk},$$

$$\lambda_m(t|\boldsymbol{b}_i, \boldsymbol{z}_i, \omega_i) = \lambda_{m0}(t)\exp\{\alpha_{m1}\eta_{i1}(t, \boldsymbol{b}_{i1}, \omega_{i1}) + \cdots + \alpha_{m4}\eta_{i4}(t, \boldsymbol{b}_{i4}, \omega_{i4}) + \boldsymbol{z}_i^T\boldsymbol{\gamma}_m\},$$

where $\boldsymbol{\omega}_i = (\omega_{i1}, \ldots, \omega_{i4})$. In this case, $\boldsymbol{\omega} = \{\omega_{11}, \ldots, \omega_{14}, \ldots, \omega_{n1}, \ldots, \omega_{n4})$ and $\boldsymbol{\omega}^0 = \mathbf{1}$ represents no perturbation. This single-case perturbation is used to detect these influential subjects, whose survival times have a low chance of occurrence given their subject-specific longitudinal profiles. Inspecting Figures 3 and 4 reveal the following influential subjects. Specifically, in Figure 3, patients 15, 28, 40, 70, 81, 96, 117, 136, 228, and 282 were detected to be the most influential for $k = 1$ and patients 15, 28, 30, 40, 70, 94, 136, 220, and 234 were detected to be the most influential for $k = 2$. Figure 4 shows that patients 15, 28, 70, 81, 136, and 150 were detected to be the most influential for $k = 3$ and patients 9, 15, 18, 28, 70, 94, 96, and 136 were detected to be the most influential for $k = 4$.

The fourth perturbation involves perturbing the prior distributions as follows:

$$\beta_k \sim N_7(\beta_k^0 + \omega_{\beta0}\mathbf{1}_7, \boldsymbol{H}_{\beta k}^0/\omega_{\beta 1}),$$
$$\alpha_m \sim N_4(\alpha_m^0 + \omega_{\alpha0}\mathbf{1}_4, \omega_{\alpha1}^{-1}\boldsymbol{H}_{\alpha m}^0),$$
$$\Sigma^{-1} \sim \text{Wishart}_4(\rho_0, \omega_{\Sigma}^{-1}\boldsymbol{R}^0),$$
$$\boldsymbol{\Phi}_k^{-1} \sim \text{Wishart}_2(\rho_k^0, \omega_{\varphi}^{-1}\boldsymbol{R}_{\varphi k}^0),$$
$$\gamma_m \sim N_6(\gamma_m^0 + \omega_{\gamma0}\mathbf{1}_6, \omega_{\gamma1}^{-1}\boldsymbol{H}_{\gamma m}^0), h_{ml} \sim \Gamma(\tau_{\lambda0}, \omega_{\lambda}\tau_{\lambda 1}).$$

In this case, $\boldsymbol{\omega} = \{\omega_{\beta0}, \omega_{\beta1}, \omega_{\alpha0}, \omega_{\alpha1}, \omega_{\gamma0}, \omega_{\gamma1}, \omega_{\lambda}, \omega_{\Sigma}, \omega_{\varphi}\}$, and $\boldsymbol{\omega}^0 = \{0, 1, 0, 1, 0, 1, 1, 1, 1\}$ represents no perturbation. This perturbation assesses the sensitivity of the Bayes factor to minor changes in the prior distributions. Based on the local influence measures $|\boldsymbol{h}_{\max}|$ for the Bayes factor and the diagonal elements of $G(\boldsymbol{\omega}^0)$, Figures 5(a) and (b) shows that perturbing the prior of h_{ml} has a large impact on the Bayesian analysis.

The last perturbation is a simultaneous perturbation. Specifically, we consider the following perturbation scheme:

$$y_{ik}(t_{ijk}, \omega_i, \omega_b) = \beta_{k0} + \beta_{k1}x_{i1} + \cdots + \beta_{k6}x_{i6} + \beta_{k7}t_{ijk} + \omega_b(b_{ik0} + b_{ik1}t_{ijk}) + \varepsilon_{ijk}/\omega_i$$
$$\triangleq \eta_{ik}(t_{ijk}, \boldsymbol{b}_{ik}, \omega_b) + \varepsilon_{ijk}/\omega_i,$$

$$\lambda_m(t|\boldsymbol{b}_i, \boldsymbol{z}_i, \omega_{\lambda}, \omega_b) = \lambda_{m0}(t)\exp\{\alpha_{m1}\eta_{i1}(t, \boldsymbol{b}_{i1}, \omega_b) + \cdots + \alpha_{m4}\eta_{i4}(t, \boldsymbol{b}_{i4}, \omega_b) + \boldsymbol{z}_i^T\boldsymbol{\gamma}_m + \omega_{\lambda}\},$$

and a subset of the h_{ml}'s and the priors of $\boldsymbol{\beta}_k$, $\boldsymbol{a}_m$, $\boldsymbol{\gamma}_m$, h_{ml}, $\boldsymbol{\Sigma}^{-1}$ and $\boldsymbol{\Phi}_k^{-1}$ are pertarbed. In this case, we have

$$\omega = \{\omega_1, \ldots, \omega_n, \omega_b, \omega_\lambda, \omega_{h1}, \ldots, \omega_{hG}, \omega_{\beta 0}, \omega_{\beta 1}, \omega_{\alpha 0}, \omega_{\alpha 1}, \omega_{\gamma 0}, \omega_{\gamma 1}, \omega_h, \omega_\Sigma, \omega_\varphi\}$$

and $\boldsymbol{\omega}^0 = \{1, \ldots, 1, 1, 0, 1, 1, \ldots, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1\}$ represents no perturbation. We are interested in examining the sensitivity of all components of the Bayesian analysis to such simultaneous perturbations. Based on all the subcomponents of $|\boldsymbol{h}_{\max}|$ and g_{ii}, Figure 5(c) identifies influential perturbations $b_{ik0} + b_{ik1}t$, $\boldsymbol{z}_i^T \boldsymbol{\gamma}_m$, h_{ml}s and the prior distribution of the h_{ml}s as well as the three most influential patients 103, 130, and 780. Finally, we deleted the three influential subjects 103, 130, and 780 and recalculated the posterior estimates of the parameters for the IBCSG data (Table 1). Inspecting Table 1 indicates that many subcomponents of $\boldsymbol{\beta}_k$ and $\boldsymbol{a}_m$ are very sensitive to the deletion of these three subjects.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Amari, SI. Lecture Notes in Statistics. 2. Vol. 28. Berlin: Springer-Verlag; 1990. Differential-Geometrical Methods in Statistics.

Arjas E, Gasbarra D. Nonparametric bayesian inference from right censored survival data, using the gibbs sampler. Statistica Sinica. 1994; 4:505–524.

Brown ER, lbrahim JG. A bayesian semiparametric joint hierarchical model for longitudinal and survival data. Biometrics. 2003a; 59:221–228. [PubMed: 12926706]

Brown ER, lbrahim JG. Bayesian approaches to joint cure rate and longitudinal models with applications to cancer vaccine trials. Biometrics. 2003b; 59:686–693. [PubMed: 14601770]

Brown ER, Ibrahim JG, DeGruttola V. A flexible B-spline model for multiple longitudinal biomarkers and survival. Biometrics. 2005; 61:64–73. [PubMed: 15737079]

Chen MH, Ibrahim JG, Sinha D. Bayesian inference for multivariate survival data with a surviving fraction. Journal of Multivariate Analysis. 2002; 80:101–126.

Chen MH, Ibrahim JG, Sinha D. A new joint model for longitudinal and survival data with a cure fraction. Journal of Multivariate Analysis. 2004; 91:18–34.

Chi YY, Ibrahim JG. Joint models for multivariate longitudinal and multivariate survival data. Biometrics. 2006; 62:432–445. [PubMed: 16918907]

Chi Y, Ibrahim JG. A New class of joint models for longitudinal and survival data accomodating zero and non-zero cure fractions: A case study of an international breast cancer study group trial. Statistica Sinica. 2007; 17:445–462.

Cook RD. Assessment of local influence (with Discussion). Journal of the Royal Statistical Society, Series B: Methodological. 1986; 48:133–169.

De Gruttola V, Tu XM. Modelling progression of cd-4 lymphocyte count and its relationship to survival time. Biometrics. 1994; 50:1003–1014. [PubMed: 7786983]

Dobson A, Henderson R. Diagnostics for joint longitudinal and dropout time modeling. Biometrics. 2003; 59:741–751. [PubMed: 14969451]

Dunson, D. Nonparametric Bayes applications to biostatistics. In: Hjort, N.; Holmes, C.; Muller, P.; Walker, S., editors. Bayesian Nonparametrics in Practice. Cambridge, UK: Cambridge University Press; 2009. p. 223-273.

Faucett C, Thomas D. Simultaneously modeling censored survival data and repeatedly measured covariates: A gibbs sampling approach. Statistics in Medicine. 1996; 15:1663–1685. [PubMed: 8858789]

Hanson T, Branscum A, Johnson W. Predictive comparison of joint longitudinal-survival modeling: A case study illustrating competing approaches (with discussion). Lifetime Data Analysis. 2011; 17:3–28. [PubMed: 20369294]

Henderson R, Diggle P, Dobson A. Joint modeling of longitudinal measurements and event time data. Biostatistics. 2000; 4:465–480. [PubMed: 12933568]

Ibrahim, JG.; Chen, MH.; Sinha, D. Bayesian Survival Analysis. Springer Series in Statistics. New York: Springer-Verlag; 2001.

Ibrahim JG, Chen MH, Sinha D. Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine studies. Statistica Sinica. 2004; 14:863–883.

Law NJ, Taylor JMG, Sandler HM. The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure. Biostatistics. 2002; 3:547–563. [PubMed: 12933597]

Pawitan Y, Self S. Modeling disease marker processes in aids. Journal of the American Statistical Association. 1993; 83:719–726.

Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. Statistics in Medicine. 2011; 30:1366–1380. [PubMed: 21337596]

Rizopoulos D, Verbeke G, Molenberghs G. Shared parameter models under random effects misspecification. Biometrika. 2008; 95:63–74.

Sinha D. Semiparametric Bayesian analysis of multiple event time data. Journal of the American Statistical Association. 1993; 88:979–983.

Song X, Davidian M, Tsiatis AA. A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. Biometrics. 2002; 58:742–753. [PubMed: 12495128]

Tsiatis A, Davidian M. An overview of joint modeling of longitudinal and time-to-event data. Statistica Sinica. 2004; 14:793–818.

Tsiatis AA, DeGruttola V, Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and cd4 counts in patients with aids. Journal of the American Statistical Association. 1995; 90:27–37.

Wang Y, Taylor JMG. Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. Journal of the American Statistical Association. 2001; 96:895–905.

Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics. 1997; 53:330–339. [PubMed: 9147598]

Xu J, Zeger SL. The evaluation of multiple surrogate endpoints. Biometrics. 2001; 57:81–87. [PubMed: 11252622]

Yu M, Law NJ, Taylor JMG, Sandler HM. Joint longitudinal-survival-cure models and their application to prostate cancer. Statistica Sinica. 2004; 14:835–862.

Zhu H, Ibrahim J, Lee SY, Zhang H. Perturbation selection and influence measures in local influence analysis. Annals of Statistics. 2007; 35:2565–2588.

Zhu H, Ibrahim JG, Tang NS. Bayesian local influence analysis: a geometric approach. Biometrika. 2011; 98:307–323.

Zhu H, Lee SY. Local influence for incomplete-data models. Journal of the Royal Statistical Society, Series B: Statistical Methodology. 2001; 63:111–126.

Figure 1.
Results of single-case perturbation to longitudinal profiles for the IBCSG data set: Index plots of the local influence measure $|\boldsymbol{h}_{\max}|$ corresponding to $k = 1$ (left upper panel), $k = 2$ (right upper panel), $k = 3$ (left lower panel), and $k = 4$ (right lower panel).
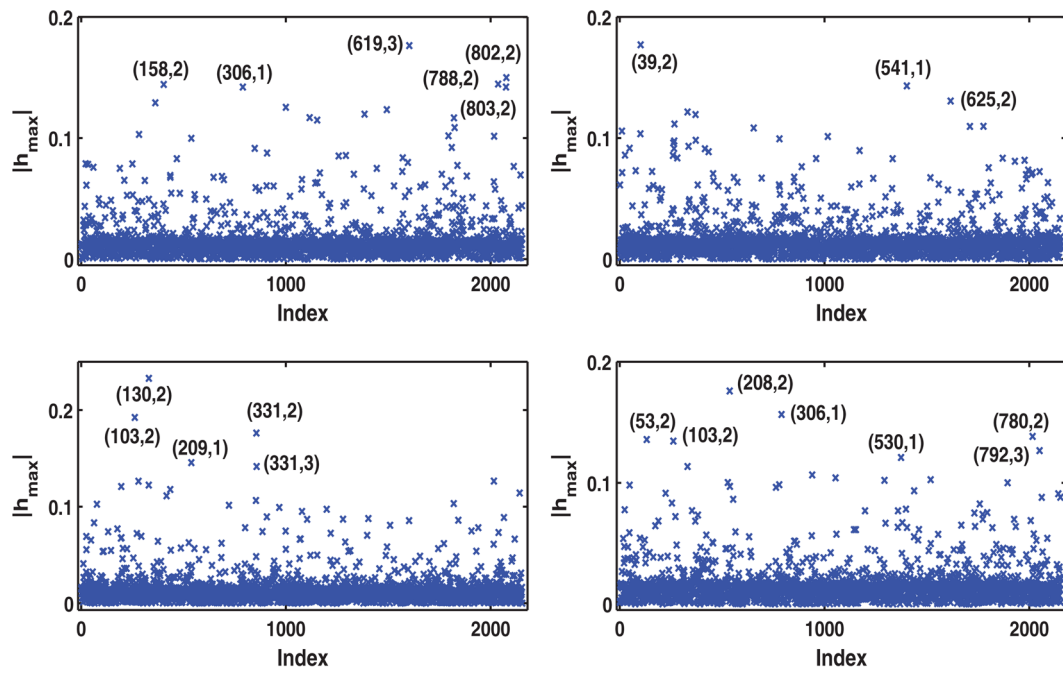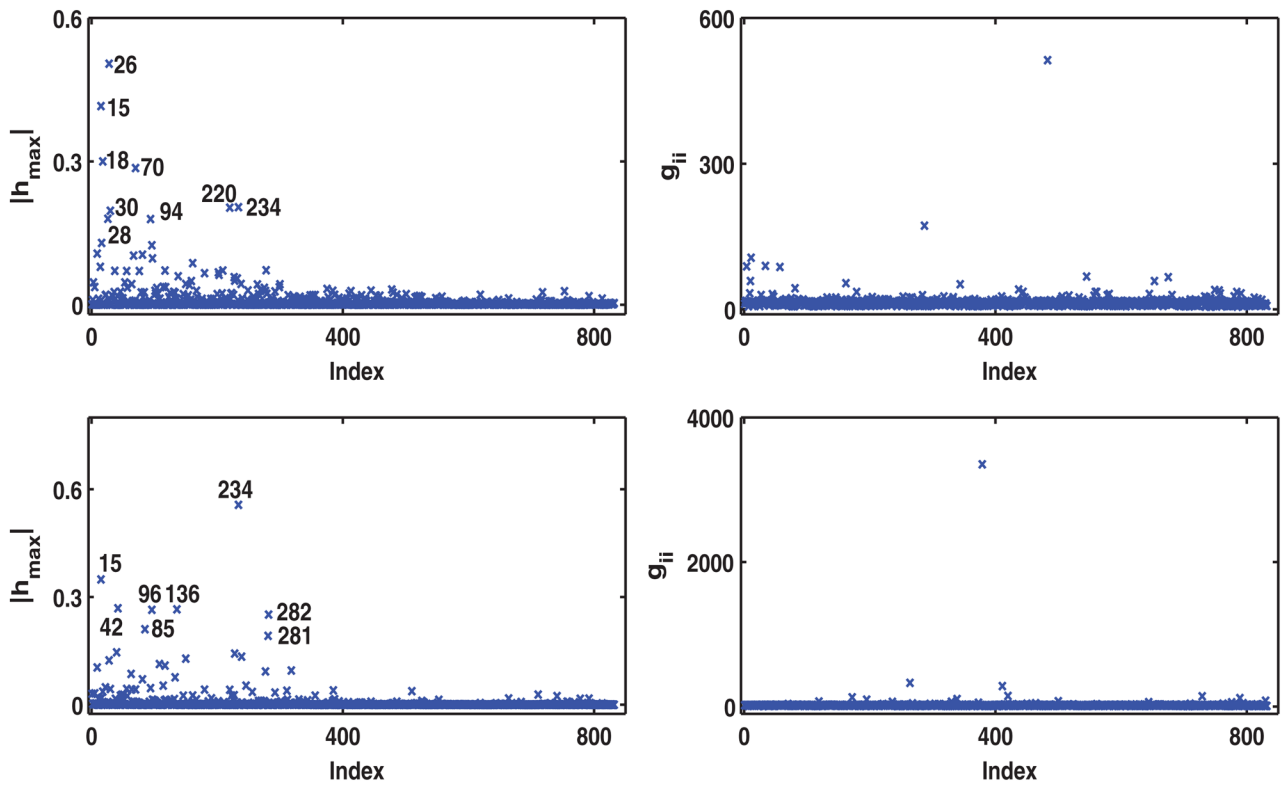
Figure 2.
Results of single-case perturbation to marginal hazard functions for the IBCSG data set: Index plots of the local influence measure h_{max} for ω_{ins} with $m = 1$ (left upper panel), and for ω_{ins} with $m = 2$ (left lower panel), and g_{ii} with $m = 1$ (right upper panel), and $m = 2$ (right lower panel).

Figure 3.
Results of perturbing the shared random effects for the IBCSG data set: Index plots of the local influence measure $|\boldsymbol{h}_{\max}|$ for perturbing b_{ik} with $k = 1$ (left upper panel), and $k = 2$ (left lower panel), g_{ii} with $k = 1$ (right upper panel) and $k = 2$ (right lower panel).
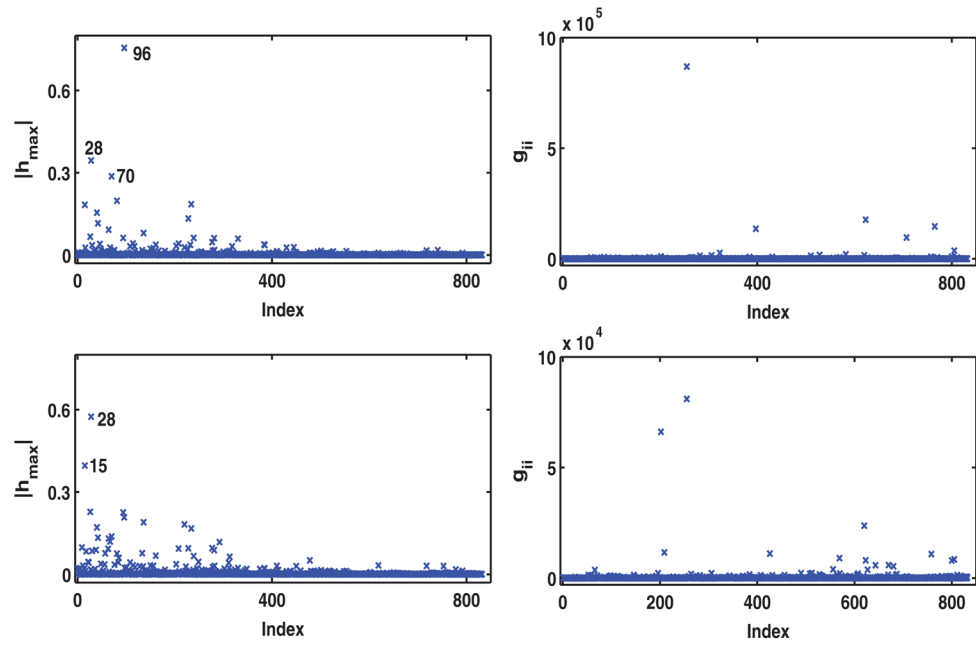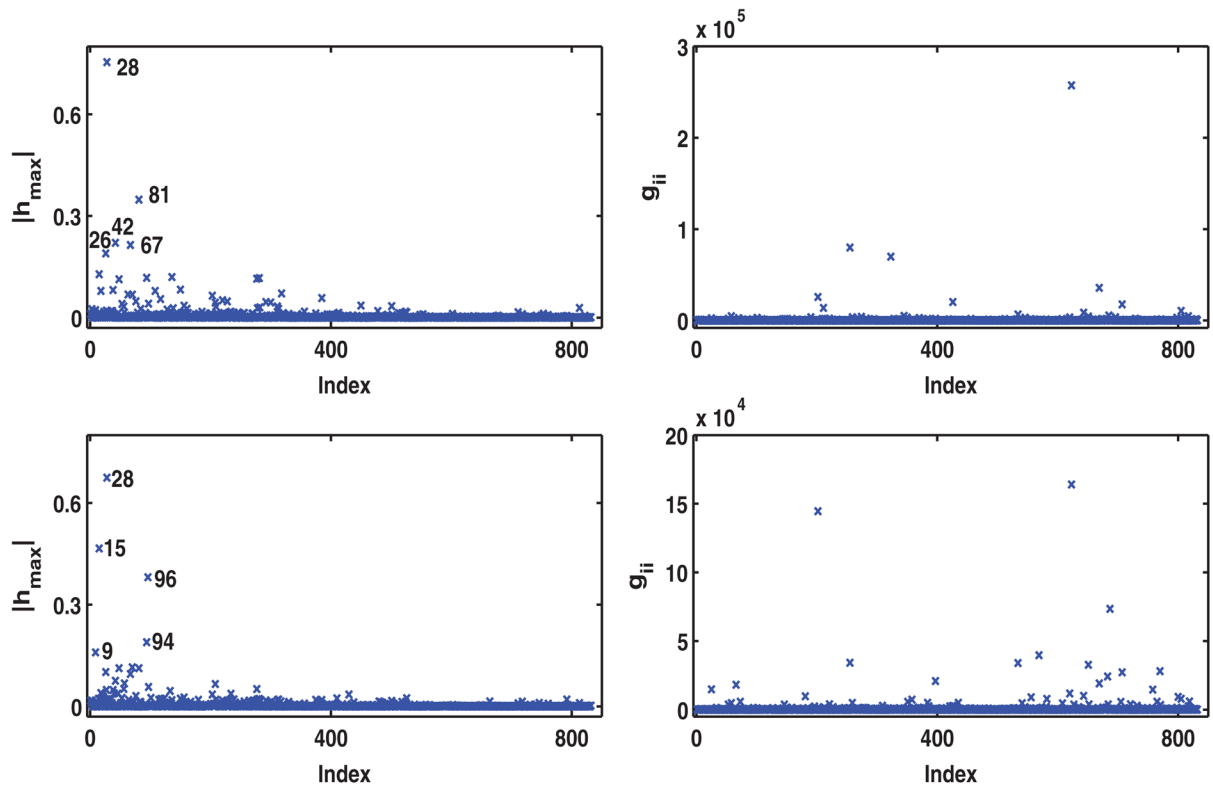
Figure 4.
Results of perturbing the shared random effects for the IBCSG data set: Index plots of the local influence measure $|\boldsymbol{h}_{\max}|$ for perturbing b_{ik} with $k = 3$ (left upper panel), and $k = 4$ (left lower panel), g_{ii} with $k = 3$ (right upper panel) and $k = 4$ (right lower panel).
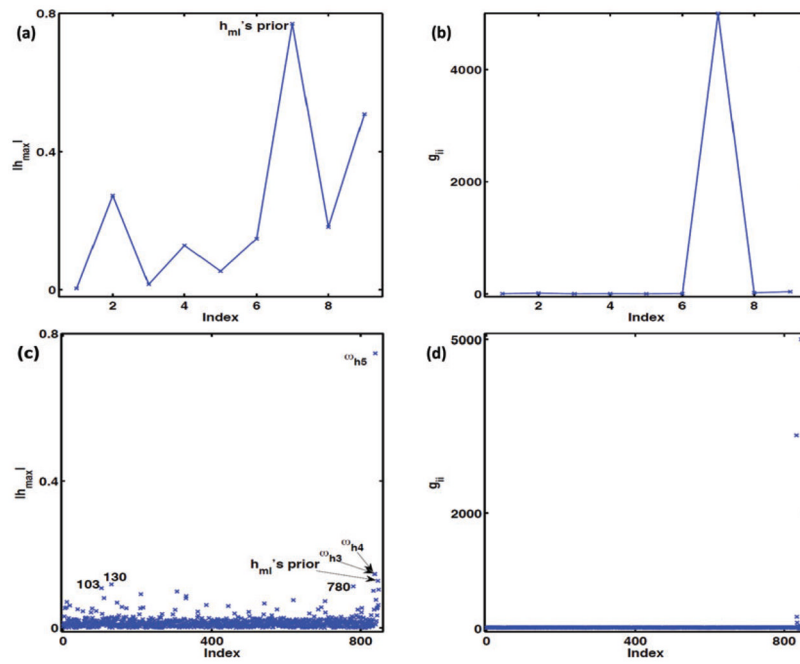
Figure 5.
Results of the IBCSG data set. Perturbing the prior distributions: (a) index plot of the local influence measure $|h_{\max}|$ (the left panel); and (b) index plot of g_{ii} (the right panel) for perturbing priors of the parameters $\boldsymbol{\beta}_k$, $\boldsymbol{a}_m$, $\boldsymbol{\gamma}_m$, $h_{m\,l}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}_1$, ..., $\boldsymbol{\Phi}_4$. The simultaneous perturbation: (c) index plot of the local influence measure $|h_{\max}|$ (the left panel); and (d) index plot of g_{ii} (the right panel).

Table 1

The IBCSG data set: Posterior means (Mean) of the parameters and their posterior standard deviations (SD) with and without influential patients 103, 130, and 780.

Par.	With Mean	With SD	Without Mean	Without SD	Par.	With Mean	With SD	Without Mean	Without SD
α_{11}	0.263	0.055	0.725	0.141	β_{10}	3.481	0.147	4.285	0.052
α_{12}	−0.838	0.042	0.090	0.142	β_{11}	−0.051	0.154	−0.265	0.082
α_{13}	−1.081	0.061	0.077	0.129	β_{12}	−0.090	0.162	−0.106	0.097
α_{14}	1.245	0.076	−1.655	0.215	β_{13}	−0.050	0.222	0.349	0.086
α_{21}	0.642	0.096	1.132	0.146	β_{14}	0.585	0.127	−0.150	0.052
α_{22}	−1.363	0.067	0.062	0.179	β_{15}	0.140	0.182	−0.320	0.059
α_{23}	−1.525	0.056	0.069	0.156	β_{16}	0.174	0.141	−0.154	0.091
α_{24}	1.650	0.121	−2.276	0.334	β_{17}	−0.735	0.056	−0.903	0.019
γ_{11}	0.047	0.240	0.062	0.214	β_{20}	5.177	0.604	5.806	0.140
γ_{12}	−0.437	0.235	−0.398	0.201	β_{21}	0.268	0.124	−0.195	0.154
γ_{13}	−0.149	0.346	0.084	0.399	β_{22}	0.152	0.200	−0.197	0.153
γ_{14}	−1.185	0.243	−0.250	0.164	β_{23}	−0.329	0.219	0.440	0.205
γ_{15}	1.304	0.166	1.058	0.159	β_{24}	0.344	0.246	0.140	0.158
γ_{16}	−0.790	0.212	−0.322	0.154	β_{25}	0.298	0.135	0.204	0.130
γ_{21}	−1.280	0.480	−1.044	0.367	β_{26}	0.377	0.254	0.012	0.122
γ_{22}	−1.851	0.449	−1.685	0.377	β_{27}	−0.863	0.112	−1.014	0.048
γ_{23}	1.525	0.682	1.695	0.723	β_{30}	4.289	0.162	5.048	0.120
γ_{24}	−2.911	0.542	−1.317	0.300	β_{31}	0.347	0.094	0.059	0.118
γ_{25}	2.129	0.245	1.589	0.334	β_{32}	0.213	0.136	0.158	0.125
γ_{26}	−2.929	0.430	−1.726	0.239	β_{33}	−0.428	0.205	−0.089	0.180
σ_{11}	3.941	0.165	4.259	0.205	β_{34}	0.594	0.120	0.148	0.109
σ_{12}	1.588	0.115	1.642	0.120	β_{35}	0.268	0.161	−0.061	0.119
σ_{13}	2.444	0.125	2.524	0.131	β_{36}	0.820	0.104	0.159	0.098
σ_{14}	2.537	0.122	2.533	0.130	β_{37}	−0.823	0.037	−0.855	0.057
σ_{22}	4.163	0.173	3.660	0.164	β_{40}	3.974	0.623	4.560	0.071

Par.	With Mean	SD	Without Mean	SD	Par.	With Mean	SD	Without Mean	SD
σ_{23}	2.716	0.128	2.386	0.129	β_{41}	0.047	0.192	0.017	0.058
σ_{24}	2.112	0.131	2.322	0.129	β_{42}	−0.182	0.241	−0.029	0.065
σ_{33}	4.672	0.171	4.449	0.170	β_{43}	−0.019	0.293	0.130	0.076
σ_{34}	3.129	0.137	3.395	0.142	β_{44}	0.853	0.290	0.215	0.036
σ_{44}	4.695	0.173	5.243	0.177	β_{45}	0.151	0.102	−0.170	0.055
$\varphi_{1,11}$	1.516	0.179	1.073	0.207	β_{46}	0.724	0.140	0.158	0.028
$\varphi_{1,12}$	−0.567	0.110	−0.356	0.098	β_{47}	−0.655	0.115	−0.661	0.011
$\varphi_{1,22}$	0.479	0.078	0.380	0.074	$\varphi_{3,11}$	0.685	0.081	1.067	0.117
$\varphi_{2,11}$	1.163	0.133	1.539	0.156	$\varphi_{3,12}$	−0.267	0.056	−0.376	0.076
$\varphi_{2,12}$	−0.241	0.079	−0.289	0.121	$\varphi_{3,22}$	0.626	0.079	0.417	0.068
$\varphi_{2,22}$	0.654	0.084	0.644	0.139					
$\varphi_{4,11}$	0.537	0.088	0.328	0.072					
$\varphi_{4,12}$	−0.144	0.056	−0.090	0.053					
$\varphi_{4,22}$	0.403	0.057	0.404	0.057					