# Parametric joint modelling of longitudinal and survival data

**Diana C. Franco-Soto [1], Antonio C. Pedroso-de-Lima [2], and Julio M. Singer [2]**

[1] Departamento de Estadística, Universidad Nacional de Colombia, Bogotá, Colombia

[2] Departmento de Estatística, Universidade de São Paulo, São Paulo, Brazil

---

**Address for correspondence:** Antonio Carlos Pedroso-de-Lima, Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, São Paulo - SP, 05508-090, Brazil.

**E-mail:** acarlos@ime.usp.br.

**Phone:** (+55) 11 3091 6232.

**Fax:** (+55) 11 3091 6130.

---

**Abstract:** We consider parametric joint modelling of longitudinal measurements and survival times, motivated by a study conducted at the Heart Institute (Incor), São Paulo, Brazil, with the objective of evaluating the impact of B-type Natriuretic Peptide (BNP) collected at different instants on the survival of patients with congestive heart failure (CHF). We employ a linear mixed model for the longitudinal response and a Birnbaum-Saunders model for the survival times allowing the inclusion of subjects without longitudinal observations. We derive maximum likelihood

estimators of the model parameters and consider their statistical properties. We also conduct a simulation study to compare the true survival probabilities with dynamic predictions obtained from the fit of the proposed joint model and to evaluate the robustness of the method for estimating the parameters with respect to misspecification of the parametric distribution of the survival response. Finally, the proposed joint model is applied to the cohort of 1609 patients with CHF, of which 1080 have no BNP measurements. The parameter estimates and their standard errors obtained via i) the traditional approach, where only individuals with at least one measurement of the longitudinal response are included and ii) the proposed approach, which includes survival information from all individuals are compared with those obtained via marginal (longitudinal and survival) models. The results suggest that an increase in the number of subjects with measurements of the longitudinal response lead to an increase in the precision of parameter estimates, including those related to the association between the longitudinal and survival responses as well as an improvement in the quality of survival predictions.

---

**Key words:** Birnbaum-Saunders model; linear mixed model; repeated measurements

# 1 Introduction

In many studies, repeated measurements of one or more variables (longitudinal responses), time until the occurrence of one or more events (survival responses) and additional observations on explanatory variables are collected on a set of subjects in

order to characterize their relationship. This is the case of a study conducted at the Heart Institute (Incor), São Paulo, Brazil, where data related to i) longitudinal measurements of B-type Natriuretic Peptide (BNP) levels (pg/mL), ii) the time between admission to the study and the date of death or censoring (in months) as well as iii) the values of basal covariates were collected on patients with Congestive Heart Failure (CHF) to identify prognostic factors for the time to death.

In practical situations, data with this nature are often analyzed considering the longitudinal and survival responses separately as noted in Rizopoulos (2010), among others. However, there are two scenarios in which it is more appropriate to perform a joint modelling: when interest is to analyze the behavior of the longitudinal response, considering a possible dependence of time to dropout, potentially informative, treated as the survival response [Hogan and Laird (1997ab), Diggle, Farewell and Henderson (2007)] and when interest is to analyze the time-to-event considering the effect of the longitudinal response measurements [Wulfsohn and Tsiatis (1997), Henderson, Diggle, Dobson (2000), Rizopoulos (2010), Crowther, Abrams and Lambert (2012)]. Different authors suggest that in these cases joint modelling can facilitate the understanding of the mechanisms underlying the phenomenon under investigation and can improve the properties of parameter estimators, being an appealing alternative that has attracted the attention of recent research [Tsiatis and Davidian (2004), Yu et al. (2004), Diggle, Sousa and Chetwynd (2008), Wu et al. (2012), Rizopoulos (2012a) or Gould et al. (2014)].

A naive approach when the interest lies exclusively in the survival component, is to consider the longitudinal response as a time-dependent covariate in the Cox model, which requires that the time-dependent covariate values be known exactly at each

instant of failure and further, that the time-dependent covariates are external, as described by Kalbfleisch and Prentice (2002). This approach may not be appropriate because the longitudinal observations are usually measured intermittently and subject to errors. Additionally, it may be influenced by the occurrence of the event under investigation [Hu, Tsiatis and Davidian (1998), Greene and Cai (2004), Rizopoulos (2010)]. An alternative is the two-stage approach, where a model for longitudinal response is initially fitted to the data and, using the values of the first stage estimated individual longitudinal, are subsequently incorporated as a time-dependent covariate in the Cox model [Tsiatis, DeGruttola and Wulfsohn (1995), Yu, Lin and Taylor (2008), Albert and Shih (2010)]. Despite the advantage of its simple computational implementation, this method has the limitation of not considering the effect of the survival response on the modelling of the longitudinal data. Another alternative is to estimate the model parameters by maximizing the likelihood function corresponding to the joint distribution of longitudinal and survival responses [Wulfsohn and Tsiatis (1997), Henderson, Diggle and Dobson (2000), Hsieh, Tseng and Wang (2006), Crowther, Abrams and Lambert (2012), Rizopoulos (2010, 2012a)]. Although the computational implementation of this approach is more complex, it has the advantage of using longitudinal and survival data simultaneously in the process of estimating model parameters. This approach is adopted in this paper.

In this context, most authors consider the Cox model to describe survival times [see Wulfsohn and Tsiatis (1997), Henderson, Diggle and Dobson (2000), Rizopoulos (2010, 2012a), among others] although log-normal and Weibull parametric models have also been considered for such purposes in Schluchter (1992), Pawitan and Self (1993), DeGruttola and Tu (1994). Linear mixed models are commonly employed to represent the longitudinal component. The usual methods, however, only use data for

subjects that have at least one measurement of the longitudinal response. In studies where such measurements are not recorded for some participants, the corresponding estimates may be biased or less efficient.

We propose a joint modelling methodology that incorporates the survival data of subjects without measurements of the longitudinal response using linear mixed models to describe such response and Birnbaum-Saunders models to describe the survival response. Birnbaum-Saunders distributions seem appropriate in the context of CHF because in chronic cardiac diseases, a cumulative damage caused by several risk factors may lead to a degradation and to a consequent failure, an inherent feature of such models, as described in Galea, Leiva and Paula (2004), Leiva et al. (2007), Barros, Paula and Leiva (2008), Balakrishnan et al. (2009) or Leiva et al. (2011).

In Section 2, we present the model and discuss inferential aspects including the dynamic prediction of the survival probability based on the available data up to the instant when we want to make the prediction. In Section 3, we summarize the results of a simulation study designed to compare the true survival probabilities with dynamic predictions obtained from the fitted model and to evaluate the robustness of the method for estimating model parameters with respect to the misspecification of the time-to-event distribution. In Section 4 we analyze the data that motivated our research. We conclude with a discussion and suggestions for future research in Section 5.

# 2 Methodology

Consider a set of $n$ subjects followed over the time interval $[0, \tau)$, $\tau > 0$, and suppose that for the $i$-th subject ($i = 1, \ldots, n$) we observe: i) a sequence of measurements of a longitudinal response $y_{ij} = \{y_i(t_{ij}), j = 1, \ldots, n_i\}$ summarized in $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^\top$ and occurring at times $t_{ij}$ represented in $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})^\top$, ii) a record of the time between admission to the study and the occurrence of the event of interest ($T_i$) or censoring ($C_i$), summarized in $Z_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$, where $I(\cdot)$ denotes the indicator function and iii ) values of $p_h$ explanatory variables expressed as $\boldsymbol{x}_{hi}(t) = [1, x_{hi1}, \ldots, x_{hia_h}, x_{hia_h+1}(t), \ldots, x_{hip_h}(t)]^\top$, the first $a_h$ of which are independent of time. The subscript $h = 1, 2$ indicates whether they correspond to the longitudinal or to the survival components, respectively.

The longitudinal response for the $i$-th subject at time $t \geq 0$ is modelled as

$$y_i(t) = m_i(t) + e_i(t) \tag{2.1}$$

where $m_i(t) = \mu_{1i}(t) + w_{1i}(t)$ denotes the true value of the longitudinal response, specified as a function of a mean response $\mu_{1i}(t) = \boldsymbol{x}_{1i}(t)^\top \boldsymbol{\beta}_1$, with $\boldsymbol{\beta}_1$ representing the fixed effects corresponding to $p_1$ explanatory variables in $\boldsymbol{x}_{1i}(t)$ and the process $w_{1i}(t)$, characterized in terms of a specific time invariant random intercept for the $i$-th subject, $b_{0i} \sim N(0, \sigma_0^2)$ and $e_i(t) \sim N(0, \sigma_e^2)$ denotes the measurement error, considered independent of $b_{0i}$, for all $t \geq 0$.

We assume that the survival or censoring time observed for the $i$-th subject follows

the log-linear Birnbaum-Saunders regression model

$$V_i = \log(Z_i) = \boldsymbol{x}_{2i}^{\top}\boldsymbol{\beta}_2 + \varepsilon_i, \tag{2.2}$$

where $\boldsymbol{\beta}_2$ contains the fixed effects corresponding to $p_2$ explanatory variables in $\boldsymbol{x}_{2i}$ and the model errors $\varepsilon_i \sim \text{SinhN}(\alpha, 0, 2)$ with $\text{SinhN}(\alpha, \psi, \sigma)$ denoting the Normal hyperbolic sine distribution with $\alpha$, $\psi$ and $\sigma$ representing the shape, location and scale parameters, respectively. The associated density and survival functions are, respectively

$$f_V(v) = \left(\frac{2}{\alpha\sigma\sqrt{2\pi}}\right)\cosh\left(\frac{v-\psi}{\sigma}\right)\exp\left\{-\frac{2}{\alpha^2}\sinh^2\left(\frac{v-\psi}{\sigma}\right)\right\}, \quad v \in \Re. \tag{2.3}$$

and

$$S_V(v) = 1 - \Phi\left[\frac{2}{\alpha}\sinh\left(\frac{v-\psi}{\sigma}\right)\right], \quad v \in \Re. \tag{2.4}$$

For details on the relation between the Birnbaum-Saunders and the SinhN distributions, see Leiva et al. (2007).

In this set-up we develop the likelihood function corresponding to the joint distribution of longitudinal and survival responses. The random effects $b_{0i}$ $(i = 1, \ldots, n)$ take into account both the association between the longitudinal and survival responses and the correlation between the longitudinal observations. Considering that the censoring mechanism and the observation times of the longitudinal response are not informative and assuming independence between survival and censoring times [see Rizopoulos

(2012a)], the joint likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\left[\int p(v_i,\delta_i|b_{0i};\boldsymbol{\theta}_z,\boldsymbol{\beta}_1)\left[\prod_{j=1}^{n_i} p(y_i(t_{ij})|b_{0i};\boldsymbol{\theta}_y)\right]p(b_{0i};\theta_b)db_{0i}\right]^{\omega_i} \times [p(v_i,\delta_i;\boldsymbol{\theta}_{z_0})]^{(1-\omega_i)}, \tag{2.5}$$

where $\omega_i = 1$ if the $i$-th subject has at least one observation of the longitudinal response and $\omega_i = 0$, otherwise, $\boldsymbol{\theta} = (\boldsymbol{\theta}_z^\top, \boldsymbol{\theta}_y^\top, \theta_b)^\top$, with $\boldsymbol{\theta}_z = (\boldsymbol{\theta}_{z0}^\top, \gamma)^\top$ and $\boldsymbol{\theta}_{z0} = (\alpha, \boldsymbol{\beta}_2^\top)^\top$ denoting the vectors containing the parameters for the survival responses for subsets of subjects that have or have not measurements of the longitudinal outcome, respectively, $\boldsymbol{\theta}_y = (\boldsymbol{\beta}_1^\top, \sigma_e^2)^\top$ denotes the vector containing the longitudinal response parameters and $\theta_b = \sigma_0^2$. Additionally, letting $M_i(t) = \{m_i(u), 0 \leq u \leq t\}$ denote the history of the true unobserved longitudinal process up to time $t$, we assume that

$$p(v_i,\delta_i|b_{0i};\boldsymbol{\theta}_z,\boldsymbol{\beta}_1) = [f_i(v_i|M_i(v_i);\boldsymbol{\theta}_z,\boldsymbol{\beta}_1)]^{\delta_i}[S_i(v_i|M_i(v_i);\boldsymbol{\theta}_z,\boldsymbol{\beta}_1)]^{(1-\delta_i)}, \tag{2.6}$$

where $f_i(\cdot)$ and $S_i(\cdot)$ respectively denote the probability density and survival functions of the SinhN distribution with shape parameter $\alpha > 0$, scale parameter $\sigma = 2$ and location parameter

$$\psi_{\gamma i} = \boldsymbol{x}_{2i}^\top\boldsymbol{\beta}_2 + \gamma m_i(v_i) = \boldsymbol{x}_{2i}^\top\boldsymbol{\beta}_2 + \gamma[\boldsymbol{x}_{1i}(v_i)^\top\boldsymbol{\beta}_1 + b_{0i}], \tag{2.7}$$

$p(y_i(t_{ij})|b_{0i};\boldsymbol{\theta}_y)$ and $p(b_{0i};\theta_b)$ representing Normal probability density functions for the longitudinal response and the random effects, respectively. The parameter $\gamma$ measures the association between the longitudinal and survival processes. Furthermore, we assume that

$$p(v_i,\delta_i;\boldsymbol{\theta}_{z0}) = [f_i(v_i;\boldsymbol{\theta}_{z0})]^{\delta_i}[S_i(v_i;\boldsymbol{\theta}_{z0})]^{(1-\delta_i)}, \tag{2.8}$$

where $f_i(\cdot)$ and $S_i(\cdot)$ are as in (2.6) with location parameter $\psi_{0i} = \boldsymbol{x}_{2i}^\top \boldsymbol{\beta}_2$.

Explicit expressions for the terms composing the likelihood function (2.5) are given in the Appendix.

Maximum likelihood (ML) estimates of the model parameters are obtained by direct maximization of (2.5) using a quasi-Newton algorithm implemented via the `PORT` routines in the `R` optimizer `nlminb` [see Gay (1990)], one of the algorithms employed in the `JM` library (Rizopoulos, 2012b). Numerical integration is required because the integrals with respect to the random effects $b_{0i}$, $i = 1, \ldots, n$ in (2.5) have no analytical solution. For such purposes, we use Gauss-Hermite quadrature methods as suggested by Wulfsohn and Tsiatis (1997), Henderson, Diggle and Dobson (2000) or Rizopoulos (2010), among others, for situations where the random effects vector for each subject has low dimension. In particular, this is the default method employed in the `JM` library (Rizopoulos, 2012b), except in cases where a large number of random effects per subject is available. In such cases, Rizopoulos, Verbeke and Lesaffre (2009) recommend Laplace approximations. Explicit expressions for the elements of the ML estimating equations $\mathbf{U}(\boldsymbol{\theta}) = \partial(\log L(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = \mathbf{0}$ may be obtained in Franco-Soto (2014, Appendix D).

Confidence intervals and tests of hypotheses about the parameters of interest are based on empirical large sample results as suggested by Rizopoulos (2012a).

An additional interest is to predict the survival probabilities for a new subject with longitudinal measurements $Y_i(t) = \{y_i(s); 0 \leq s \leq t\}$ and values of baseline covariates contained in the vector $\boldsymbol{x}_i$ based on the fit of a joint model to a random sample $D_n = \{Z_i, \delta_i, \boldsymbol{y}_i; i = 1, \ldots, n\}$. In other words, interest lies in the conditional probability

of surviving time $u > t$, given survival up to time $t$, $\pi_i(u|t) = P(T_i \geq u|T_i > t, Y_i(t), \boldsymbol{x}_i, D_n; \boldsymbol{\theta})$. For such purpose we consider the estimator proposed by Rizopoulos (2011), namely

$$\widetilde{\pi}_i(u|t) = \frac{S_i\{u|M_i(u, \widetilde{\boldsymbol{b}}_i^{(t)}, \widehat{\boldsymbol{\theta}}); \widehat{\boldsymbol{\theta}}\}}{S_i\{t|M_i(t, \widetilde{\boldsymbol{b}}_i^{(t)}, \widehat{\boldsymbol{\theta}}); \widehat{\boldsymbol{\theta}}\}}, \tag{2.9}$$

where $\widehat{\boldsymbol{\theta}}$ corresponds the maximum likelihood estimates of $\boldsymbol{\theta}$ and $\widetilde{\boldsymbol{b}}_i^{(t)}$ denotes the mode of $\log p(\boldsymbol{b}|T_i > t, Y_i(t); \widehat{\boldsymbol{\theta}})$, where

$$\begin{aligned} p(\boldsymbol{b}|T_i > t, Y_i(t); \widehat{\boldsymbol{\theta}}) &= \frac{P(T_i > t, Y_i(t), \boldsymbol{b}; \widehat{\boldsymbol{\theta}})}{P(T_i > t, Y_i(t); \widehat{\boldsymbol{\theta}})} \\ &= \frac{P(T_i > t|\boldsymbol{b}; \widehat{\boldsymbol{\theta}}) p(Y_i(t)|\boldsymbol{b}; \widehat{\boldsymbol{\theta}}) p(\boldsymbol{b}; \widehat{\boldsymbol{\theta}})}{\int P(T_i > t|\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}) p(Y_i(t)|\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}) p(\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}) d\boldsymbol{b}_i}, \end{aligned}$$

$P\{T_i > t|\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}\} = S_i\{t|M_i(t, \boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}); \widehat{\boldsymbol{\theta}}\}$ is the survival function and

$$p(Y_i(t)|\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}) = \prod_{j=1}^{n_i(t)} p[y_i(t_{ij})|\boldsymbol{b}_i; \widehat{\boldsymbol{\theta}}],$$

with $n_i(t)$ denoting the number of longitudinal measurements of the $i$-th unit up to time $t$. The performance of (2.9) for finite samples depends on the quality of the ML estimates of $\boldsymbol{\theta}$ and on the prediction of the random effects $\boldsymbol{b}_i$.

The methodology proposed in this paper was fully implemented in R (R Development Core Team, 2013). The codes may be obtained from the authors.

# 3 Simulation

We conducted an extensive simulation study to compare the true survival probabilities with dynamic predictions based on the longitudinal data collected up to different

instants and to evaluate the robustness of the method for estimating the parameters of the proposed joint model with respect to the misspecification of the parametric distribution of the survival response. Additionally, we evaluated the performance of estimators in terms of variance and mean squared errors when the fitted model was the same as the one used to generate the data.

The longitudinal response for the $i$-th subject at time $t \geq 0$, $y_i(t)$ was generated by (2.1) under Normal distributions for both the random effects and error terms considering observation time and dichotomized CHF etiology (Chagas disease = 0 or Other cardiomyopathies = 1) as covariates. The survival response was generated either by a log-linear Birnbaum-Saunders regression model (2.2) or by a log-linear Weibull regression model including only CHF etiology as a covariate. To predict the survival probabilities, we considered only individuals with at least one measurement of the longitudinal response, *i.e.*, those for which $\omega_i = 1$ in (2.5). For the log-linear Weibull survival model, the functions $f_i(\cdot)$ and $S_i(\cdot)$ in (2.6) correspond to the probability density and survival functions of the Extreme Value distribution [see Kalbfleisch and Prentice (2002)] with scale parameter $\varsigma > 0$ and location parameter (2.7). In the Birnbaum-Saunders case, the location parameter can be expressed as

$$\psi_{\gamma_i} = \beta_{20} + \beta_{21} CHF_i + \gamma[\beta_{10} + \beta_{11} v_i + \beta_{12} CHF_i + b_{0i}] \tag{3.1}$$

and the variance components are $\sigma_0^2$ and $\sigma_e^2$.

The parameter values were taken as the estimates obtained by fitting the joint models to the 529 patients with at least one longitudinal observation in the Incor data.

First we generated the longitudinal observations for each patient and then considered

the corresponding mean observation as well as the CHF etiology covariate to generate the survival data, inducing an association between the two types of response. To mimic the set-up in the Incor data, the survival times were right censored either by specifying a Type I censoring scheme with a maximum follow-up time of $\tau = 180$ or by randomly selecting censoring times from a Uniform distribution in the $[0, \tau]$ interval.

Thirty two different scenarios resulting for the combination of the 2 time-to-event distributions (Birnbaum-Saunders or Weibull), 4 sample sizes ($n = 100, 250, 500, 1000$) and 4 censoring percentages ($p_c = 0\%, 25\%, 50\%, 75\%$) were considered. For each scenario we generated 500 samples and, for each sample, we randomly selected 95% of the subjects to fit both the joint models (the first with a Birnbaum-Saunders time-to-event component and the second with a Weibull component); the remaining 5% were considered to estimate the conditional survival probabilities based on the estimator $\widetilde{\pi}_i(t + \Delta t|t)$ in (2.9), as a function of the longitudinal response observation times $t = 0, 24, 48, 72, 96$ and the time increments $\Delta t = 12, 24, 36$. The true probabilities $\pi_i(t+\Delta t|t)$ were computed using the true parameter and the true (generated) random effects values. The comparison was carried out in terms of the absolute differences between estimated conditional survival probabilities and their true values. For the scenarios where the survival model considered to generate the data was the same as the one used to fit the data, the performance of estimators was evaluated via the bias $(\overline{\widehat{\theta}} - \theta)$ with $\overline{\widehat{\theta}} = 500^{-1} \sum_{l=1}^{500} \widehat{\theta}_l$, the relative bias $[(\overline{\widehat{\theta}} - \theta)/\theta]$ as well as via the square root of the mean squared error (MSE), $[500^{-1} \sum_{l=1}^{500} (\widehat{\theta}_l - \theta)^2]^{1/2}$.

For each scenario, we constructed histograms for such differences based on the 500 samples and computed the corresponding 95-th percentile for each combination of $t$

and $\Delta t$.

We show the results obtained for a sample of size $n = 1000$. Results for the remaining scenarios may be obtained in

http://www.ime.usp.br/∼acarlos/Diana/ApendiceF.pdf.

In Table 1 we present these values for the case where the data were generated and fitted according to a Birnbaum-Saunders model, no random censoring and 7% Type I censoring. The number of differences corresponds to the available longitudinal measurements for the 50 subjects selected for prediction in the 500 samples.

Table 1: *95-th percentile and number of differences between true and estimated survival probabilities, Birnbaum-Saunders generated and fitted model,* $n = 1000$*, Type I censoring = 7%, random censoring = 0%*

| | 95-th percentile | | | Number |
|---|---|---|---|---|
| | $\Delta t = 12$ | $\Delta t = 24$ | $\Delta t = 36$ | of differences |
| $t = 0$ | 0.25 | 0.34 | 0.35 | 25000 |
| $t = 24$ | 0.18 | 0.24 | 0.27 | 18354 |
| $t = 48$ | 0.12 | 0.18 | 0.21 | 11736 |
| $t = 72$ | 0.09 | 0.15 | 0.18 | 7761 |
| $t = 96$ | 0.08 | 0.12 | 0.15 | 5364 |

The results indicate that for a fixed time $t$, the differences between the true and estimated probabilities increase as $\Delta t$ increases. This is justified by the increasing distance between the time up to which longitudinal data are available and the instant for which the prediction is made. Furthermore, for $\Delta t$ fixed, there is a decrease in these differences as $t$ increases. A possible explanation is that availability of more

longitudinal measurements for each subject improves the predictions of the random effects involved in the proposed estimator. To verify this, we computed 95-th percentile of the empirical distribution (based on the 500 generated samples) of the differences between the true and predicted random effects. The results for the scenario described above are displayed in Table 2 and confirm our conjecture.

Table 2: *95-th percentile and number of differences between true and estimated random effects, Birnbaum-Saunders generated and fitted model,* $n = 1000$*, Type I censoring = 7%, random censoring = 0%*

| | 95-th percentile | Number of differences |
|---|---|---|
| $t = 0$ | 1.33 | 25000 |
| $t = 24$ | 0.99 | 18354 |
| $t = 48$ | 0.84 | 11736 |
| $t = 72$ | 0.76 | 7761 |
| $t = 96$ | 0.71 | 5364 |

We also computed the mean and median of the maximum (among patients) differences between the true and estimated survival probabilities for the 500 generated samples. The results, displayed in Table 3 also confirm our previous conclusions.

Table 3: *Mean and median of the maximum (among patients) differences between the true and estimated survival probabilities, Birnbaum-Saunders generated and fitted model,* $n = 1000$*, Type I censoring = 7%, random censoring = 0%*

| | $\Delta t = 12$ | | $\Delta t = 24$ | | $\Delta t = 36$ | |
|---|---|---|---|---|---|---|
| | Mean | Median | Mean | Median | Mean | Median |
| $t = 0$ | 0.30 | 0.28 | 0.41 | 0.37 | 0.43 | 0.40 |
| $t = 24$ | 0.21 | 0.19 | 0.29 | 0.26 | 0.33 | 0.30 |
| $t = 48$ | 0.14 | 0.12 | 0.20 | 0.18 | 0.24 | 0.22 |
| $t = 72$ | 0.11 | 0.09 | 0.16 | 0.14 | 0.19 | 0.17 |
| $t = 96$ | 0.08 | 0.07 | 0.13 | 0.11 | 0.16 | 0.14 |

Moreover, regarding the dynamic prediction of the conditional survival probabilities, the methods for estimating model parameters for Birnbaum-Saunders and the Weibull models were robust with respect to the wrong specification of the survival response distribution, more evidently when more subjects have a larger number of longitudinal measurements. However, as expected, better results were obtained when the fitted model was the same as the one used to generate the survival data. In these cases, the bias, the relative bias and the square root of the MSE of the parameter estimates decrease as the sample size increases or the percentage of censoring decreases. In Table 4 we present the average mean and median of the maximum (among patients) differences between the true and estimated survival probabilities for the 500 generated samples for all combinations of $t$ and $\Delta t$ values, weighted by the number of maxima in each case.

Table 4: *Weighted mean of mean and median of the maximum (among patients) differences between the true and estimated survival probabilities, Birnbaum-Saunders generated and Birnbaum-Saunders or Weibull fitted models*

| | Percent | BS generated / BS fitted | | BS generated / Weibull fitted | |
|---|---|---|---|---|---|
| $n$ | censored | Mean | Median | Mean | Median |
| 1000 | 0% | 0.22 | 0.20 | 0.25 | 0.24 |
| | 25% | 0.20 | 0.18 | 0.22 | 0.21 |
| | 50% | 0.23 | 0.23 | 0.26 | 0.26 |
| | 75% | 0.50 | 0.49 | 0.37 | 0.37 |
| 500 | 0% | 0.22 | 0.20 | 0.21 | 0.20 |
| | 25% | 0.19 | 0.17 | 0.19 | 0.18 |
| | 50% | 0.20 | 0.20 | 0.22 | 0.22 |
| | 75% | 0.42 | 0.41 | 0.32 | 0.32 |
| 250 | 0% | 0.19 | 0.17 | 0.18 | 0.16 |
| | 25% | 0.16 | 0.14 | 0.16 | 0.15 |
| | 50% | 0.18 | 0.17 | 0.19 | 0.18 |
| | 75% | 0.35 | 0.34 | 0.28 | 0.28 |
| 100 | 0% | 0.16 | 0.13 | 0.14 | 0.12 |
| | 25% | 0.13 | 0.11 | 0.14 | 0.12 |
| | 50% | 0.14 | 0.13 | 0.15 | 0.14 |
| | 75% | 0.26 | 0.25 | 0.23 | 0.23 |

In Table 5, we present the bias, relative bias and square root of the MSE for the case where data were generated from a Birnbaum-Saunders model, $n = 1000$, no random censoring, 7% Type I censoring time and were also analyzed under a Birnbaum-Saunders model.

Table 5: *Bias, relative bias and square root of the MSE of estimators for Birnbaum-Saunders generated and fitted model, $n = 1000$, Type I censoring = 7%, random censoring = 0%*

| Parameter | True value | Simulation mean | Bias | Relative bias | Square root of MSE |
|---|---|---|---|---|---|
| $\log(\alpha)$ | -0.460 | -0.396 | 0.064 | 0.140 | 0.104 |
| $\beta_{20}$ | 6.140 | 5.480 | -0.660 | -0.108 | 0.928 |
| $\beta_{21}$ | 0.100 | 0.479 | 0.379 | 3.792 | 0.561 |
| $\gamma$ | -0.460 | -0.436 | 0.024 | 0.053 | 0.065 |
| $\beta_{10}$ | 5.800 | 5.670 | -0.130 | -0.022 | 0.218 |
| $\beta_{11}$ | 0.010 | 0.009 | -0.001 | -0.127 | 0.002 |
| $\beta_{12}$ | -1.100 | -1.073 | 0.027 | 0.024 | 0.183 |
| $\log(\sigma_0)$ | 0.450 | 0.382 | -0.068 | -0.152 | 0.091 |
| $\log(\sigma_e)$ | -0.320 | -0.276 | 0.045 | 0.139 | 0.061 |

Comparing the results from the different scenarios, we observed that the longitudinal component parameter estimates are more stable than those associated to the survival component which have a poor performance for small ($n = 100$) sample sizes. In summary, we conclude that the performance of the proposed model is better when i) the survival model used to generate and fit the data is the same; ii) the sample size is larger and iii) the right censoring is smaller. In particular when the right censoring percentage is large, the parameter estimates have a poor performance and the prediction of the random effects are not as good as those obtained when more longitudinal measurements are available. This leads to a decrease in the quality of the prediction of the dynamic survival probabilities. Therefore the Birnbaum-Saunders

joint model is recommended for situations where the underlying rationale for such model is reasonable, the sample size is large and there is little right censoring.

# 4 Analysis of the Incor data

The proposed joint model was applied to the cohort of 1609 patients with CHF, of which 1080 have no BNP measurements.

An initial analysis involved a selection of "acceptable" models for the longitudinal and for the survival data fitted separately via standard techniques as suggested by Wu et al. (2012). The 529 patients with at least one longitudinal observation were considered for the former and the complete set of 1609 patients was used for the latter. In this process, each of 24 covariates were fitted individually along with CHF etiology (as suggested by the physicians) and the significant ones were subsequently fitted simultaneously to either the longitudinal or the survival component. The separate longitudinal and survival component models were sequentially refitted with the removal of the non-significant variables or grouping levels with non-significant effects in each step. The (significant) variables in the last step were chosen to compose the joint model. Observation time, CHF etiology and left atrium diameter were used as covariates for the longitudinal component of the model whereas CHF etiology and left ventricular ejection fraction were considered as covariates for the survival response.

The longitudinal component was modelled via (2.1) with

$$\boldsymbol{x}_{1i}(t) = [1, x_{1i1}(t), x_{1i2}, x_{1i3}, x_{1i4}, x_{1i5}, x_{1i6}, x_{1i7}, x_{1i8}]^\top,$$

where $x_{1i1}(t)$ denotes the time at which the response was observed, $x_{1i2}$, $x_{1i3}$, $x_{1i4}$, $x_{1i5}$, $x_{1i6}$ are dummy variables associated to the categories of CHF etiology (dilated, ischaemic, hypertensive, alcoholic or other cardiopathies), $x_{1i7}, x_{1i8}$ are dummy variables associated to the categories of left atrium diameter (augmented or missing) and $w_{1i}(t)$ consisting of a random intercept $b_{0i} \sim N(0, \sigma_0^2)$.

The survival component was modelled via (2.2) with

$$\boldsymbol{x}_{2i} = [1, x_{2i1}, x_{2i2}, x_{2i3}, x_{2i4}, x_{2i5}, x_{2i6}, x_{2i7}, x_{2i8}]^\top,$$

where $x_{2i1}, x_{2i2}, x_{2i3}, x_{2i4}, x_{2i5}$ are dummy variables associated to the categories of CHF etiology, $x_{2i6}, x_{2i7}, x_{2i8}$ are dummy variables associated to the categories of left ventricular ejection fraction (very low, low or missing) and $\varepsilon_i \sim \text{SinhN}(\alpha, 0, 2)$.

Finally, the association between both components was imposed by (2.7).

Maximum likelihood parameter estimates and their standard errors obtained via i) the proposed joint model approach, which accommodates survival information of all individuals [likelihood given by (2.5)] and ii) the traditional joint model approach, where only individuals with at least one measurement of the longitudinal response are included [likelihood given by the first component of (2.5)] were compared with those obtained via marginal (longitudinal and survival) models in each set-up. The results are summarized in Tables 6 - 9.

In case i), although standard errors are slightly smaller under the joint model, no relevant differences between the joint and marginal models longitudinal parameter estimates were detected with the exception of the time coefficient for which non-significance is more evident under the former model ($p = 0.1100$ *versus* $p = 0.0586$).

Estimates of the survival parameters and corresponding standard errors are comparable for both models and the association parameter estimate is positive and marginally significant ($p = 0.0609$).

In case ii), the standard errors of the longitudinal parameter estimates are consistently smaller under the joint model, enhancing the significance of the time coefficient ($p < 0.0001$ *versus* $p = 0.0586$). Survival parameter estimates and corresponding standard errors obtained under the marginal and joint models are relatively different, leading to changes in the significance in some cases. The association parameter estimate is negative and highly significant ($p < 0.0001$).

Table 6: *Longitudinal component results - Proposed approach (n=1609).*

| | Marginal model | | | Joint model | | |
|---|---|---|---|---|---|---|
| Parameter | Estimate | Std error | p-value | Estimate | Std error | p-value |
| Intercept | 5.0529 | 0.2710 | <0.0001 | 5.0808 | 0.2271 | <0.0001 |
| Time | 0.0040 | 0.0021 | 0.0586 | 0.0037 | 0.0023 | 0.1100 |
| Dilated cardiopathy | -1.0525 | 0.2280 | <0.0001 | -1.0352 | 0.1561 | <0.0001 |
| Ischaemic cardiopathy | -0.9640 | 0.2450 | 0.0001 | -0.9692 | 0.1923 | <0.0001 |
| Hipertensive cardiopathy | -1.2735 | 0.2268 | <0.0001 | -1.1949 | 0.1720 | <0.0001 |
| Alcoholic cardiopathy | -0.2730 | 0.3861 | 0.4799 | -0.1960 | 0.2427 | 0.4194 |
| Other cardiopathies | -0.4967 | 0.4000 | 0.2149 | -0.4280 | 0.2408 | 0.0755 |
| Augmented left atrium diameter | 1.2543 | 0.2002 | <0.0001 | 1.2047 | 0.1404 | <0.0001 |
| Missing left atrium diameter | 1.1106 | 0.3080 | 0.0003 | 1.0539 | 0.2468 | <0.0001 |
| $\log(\sigma_0)$ | 0.3874 | - | - | 0.3747 | 0.0364 | <0.0001 |
| $\log(\sigma_e)$ | -0.3267 | - | - | -0.3273 | 0.0256 | <0.0001 |

Table 7: *Survival component results - Proposed approach (n=1609).*

| | Marginal model | | | Joint model | | |
|---|---|---|---|---|---|---|
| Parameter | Estimate | Std error | p-value | Estimate | Std error | p-value |
| $\log(\alpha)$ | 1.2506 | 0.1329 | <0.0001 | 1.2401 | 0.1355 | <0.0001 |
| Intercept | 7.1560 | 0.3460 | <0.0001 | 7.0577 | 0.3530 | <0.0001 |
| Dilated cardiopathy | -1.7624 | 0.1915 | <0.0001 | -1.6388 | 0.2011 | <0.0001 |
| Ischaemic cardiopathy | 0.0945 | 0.1829 | 0.6056 | 0.1261 | 0.1836 | 0.4920 |
| Hipertensive cardiopathy | 0.3879 | 0.1673 | 0.0204 | 0.4371 | 0.1695 | 0.0099 |
| Alcoholic cardiopathy | 0.4768 | 0.2841 | 0.0933 | 0.4898 | 0.2825 | 0.0829 |
| Other cardiopathies | 0.1453 | 0.2734 | 0.5952 | 0.1788 | 0.2735 | 0.5133 |
| Very low ejection fraction | -0.9429 | 0.2135 | <0.0001 | -1.0043 | 0.2156 | <0.0001 |
| Low ejection fraction | -0.1941 | 0.1991 | 0.3296 | -0.2238 | 0.1995 | 0.2619 |
| Missing ejection fraction | -0.3156 | 0.2626 | 0.2294 | -0.3049 | 0.2620 | 0.2445 |
| $\gamma$ | - | - | - | 0.0358 | 0.0191 | 0.0609 |

Table 8: *Longitudinal component results - Traditional approach (n=529).*

| | Marginal model | | | Joint model | | |
|---|---|---|---|---|---|---|
| Parameter | Estimate | Std error | p-value | Estimate | Std error | p-value |
| Intercept | 5.0529 | 0.2710 | <0.0001 | 4.7027 | 0.1744 | <0.0001 |
| Time | 0.0040 | 0.0021 | 0.0586 | 0.0092 | 0.0018 | <0.0001 |
| Dilated cardiopathy | -1.0525 | 0.2280 | <0.0001 | -0.8847 | 0.1283 | <0.0001 |
| Ischaemic cardiopathy | -0.9640 | 0.2450 | 0.0001 | -0.7222 | 0.1505 | <0.0001 |
| Hipertensive cardiopathy | -1.2735 | 0.2268 | <0.0001 | -1.1674 | 0.1694 | <0.0001 |
| Alcoholic cardiopathy | -0.2730 | 0.3861 | 0.4799 | -0.2799 | 0.1975 | 0.1563 |
| Other cardiopathies | -0.4967 | 0.4000 | 0.2149 | -0.1495 | 0.2116 | 0.4801 |
| Augmented left atrium diameter | 1.2543 | 0.2002 | <0.0001 | 1.2418 | 0.1359 | <0.0001 |
| Missing left atrium diameter | 1.1106 | 0.3080 | 0.0003 | 1.0863 | 0.2477 | <0.0001 |
| $\log(\sigma_0)$ | 0.3874 | - | - | 0.3882 | 0.0364 | <0.0001 |
| $\log(\sigma_e)$ | -0.3267 | - | - | -0.3266 | 0.0257 | <0.0001 |

Table 9: *Survival component results - Traditional approach (n=529).*

| | Marginal model | | | Joint model | | |
|---|---|---|---|---|---|---|
| Parameter | Estimate | Std error | p-value | Estimate | Std error | p-value |
| $\log(\alpha)$ | 0.0265 | 0.0881 | 0.7636 | -0.4753 | 0.0802 | <0.0001 |
| Intercept | 5.0137 | 0.2959 | <0.0001 | 7.1990 | 0.3427 | <0.0001 |
| Dilated cardiopathy | 0.3346 | 0.1689 | 0.0476 | -0.0518 | 0.1251 | 0.6792 |
| Ischaemic cardiopathy | 0.3241 | 0.1786 | 0.0696 | 0.1722 | 0.1346 | 0.2007 |
| Hipertensive cardiopathy | 0.6687 | 0.1726 | 0.0001 | 0.1698 | 0.1291 | 0.1883 |
| Alcoholic cardiopathy | 0.4770 | 0.2905 | 0.1006 | 0.3777 | 0.2148 | 0.0787 |
| Other cardiopathies | 0.3225 | 0.2869 | 0.2609 | 0.2170 | 0.2193 | 0.3224 |
| Very low ejection fraction | -0.6077 | 0.2614 | 0.0201 | -0.0620 | 0.2090 | 0.7668 |
| Low ejection fraction | -0.2235 | 0.2579 | 0.3860 | 0.0041 | 0.2073 | 0.9843 |
| Missing ejection fraction | -0.3338 | 0.3977 | 0.4014 | -0.1503 | 0.2962 | 0.6118 |
| $\gamma$ | - | - | - | -0.4569 | 0.0428 | <0.0001 |

# 5 Discussion

We propose a methodology for joint modelling of longitudinal and survival data, which differs from the methods proposed in the literature by considering a Birnbaum-Saunders model to describe the survival response and incorporating the survival information of subjects without observations of the longitudinal response.

The results of the simulation and practical application to the Incor data when only individuals with at least one measurement of the longitudinal response are included in the joint model, suggest that the inclusion of longitudinal measurements of an appropriate response may be employed to improve the analysis of survival data. In particular, an increase in the number of subjects with measurements of the longitudinal response can improve the evidence of the association between the longitudinal and survival responses and can lead to an increase in the precision of parameter estimates. In addition, an increase in the number of observations of the longitudinal response collected in a subject can improve the quality of the prediction of survival.

The results were not so evident in the practical application to the Incor data when survival information of all individuals is considered in the joint model, probably because of the observational nature of the study, carried out during 8 years with no fixed protocol for data collection. The large proportion (67%) of patients with no measurements of the longitudinal response may have masked the association between the two components of the joint model.

Future research is definitely needed before this approach can be routinely employed in practical problems. In particular, we mention diagnostic techniques and simulation studies to determine the effect of the proportion of units with missing longitudinal

data on the parameter estimates with special attention to the one relating the association between the longitudinal and survival processes.

## Acknowledgements

This research received financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq: project 308613/2011-2) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP: projects 2013/10528-2 and 2012/21788-2), Brazil.

## Appendix

Explicit expressions for the terms composing the likelihood function (2.5)

$$f_i(v_i|M_i(v_i);\boldsymbol{\theta}_z,\boldsymbol{\beta}_1) = \left(\frac{1}{\alpha\sqrt{2\pi}}\right)\cosh\left(\frac{v_i-\psi_{\gamma i}}{2}\right)\exp\left\{-\frac{2}{\alpha^2}\sinh^2\left(\frac{v_i-\psi_{\gamma i}}{2}\right)\right\}$$
$$= \exp\left\{\log[\kappa_{i1}(v_i)] - \frac{1}{2}\left[\log(8\pi) + \kappa_{i2}^2(v_i)\right]\right\},$$

$$S_i(v_i|M_i(v_i);\boldsymbol{\theta}_z,\boldsymbol{\beta}_1) = 1 - \Phi\left[\frac{2}{\alpha}\sinh\left(\frac{v_i-\psi_{\gamma i}}{2}\right)\right] = 1 - \Phi[\kappa_{i2}(v_i)],$$

where

$$\kappa_{i1}(v_i) = \frac{2}{\alpha}\cosh\left(\frac{v_i-\psi_{\gamma i}}{2}\right) \quad \text{and} \quad \kappa_{i2}(v_i) = \frac{2}{\alpha}\sinh\left(\frac{v_i-\psi_{\gamma i}}{2}\right),$$

$$p(y_i(t_{ij})|b_{0i};\boldsymbol{\theta}_y) = \frac{1}{(2\pi\sigma_e^2)^{1/2}}\exp\left\{-\frac{[y_i(t_{ij})-m_i(t_{ij})]^2}{2\sigma_e^2}\right\}$$
$$= \frac{1}{(2\pi\sigma_e^2)^{1/2}}\exp\left\{-\frac{[y_i(t_{ij})-(\boldsymbol{x}_{1i}(t_{ij})^\top\boldsymbol{\beta}_1+b_{0i})]^2}{2\sigma_e^2}\right\},$$

$$p(b_{0i};\theta_b) = \frac{1}{(2\pi\sigma_0^2)^{1/2}}\exp\left\{-\frac{b_{0i}^2}{2\sigma_0^2}\right\}.$$

$$f_i(v_i;\boldsymbol{\theta}_{z0}) = \left(\frac{1}{\alpha\sqrt{2\pi}}\right)\cosh\left(\frac{v_i-\psi_{0i}}{2}\right)\exp\left\{-\frac{2}{\alpha^2}\sinh^2\left(\frac{v_i-\psi_{0i}}{2}\right)\right\}$$
$$= \exp\left\{\log[\xi_{i1}(v_i)] - \frac{1}{2}\left[\log(8\pi)+\xi_{i2}^2(v_i)\right]\right\},$$

and

$$S_i(v_i;\boldsymbol{\theta}_{z0}) = 1-\Phi\left[\frac{2}{\alpha}\sinh\left(\frac{v_i-\psi_{0i}}{2}\right)\right] = 1-\Phi[\xi_{i2}(v_i)],$$

where

$$\xi_{i1}(v_i) = \frac{2}{\alpha}\cosh\left(\frac{v_i-\psi_{0i}}{2}\right) \quad \text{and} \quad \xi_{i2}(v_i) = \frac{2}{\alpha}\sinh\left(\frac{v_i-\psi_{0i}}{2}\right).$$

The Gauss-Hermite quadrature approximation of the logarithm of (2.5), namely,

$$l(\boldsymbol{\theta}) = \sum_{i=1}^{n}\omega_i\log[p(v_i,\delta_i,\boldsymbol{y}_i;\boldsymbol{\theta})] + \sum_{i=1}^{n}(1-\omega_i)\log[p(v_i,\delta_i;\boldsymbol{\theta}_{z_0})]$$
$$= \sum_{i=1}^{n}\omega_i\log\left[\int p(v_i,\delta_i|b_{0i};\boldsymbol{\theta}_z,\boldsymbol{\beta}_1)\left\{\prod_{j=1}^{n_i}p[y_i(t_{ij})|b_{0i};\boldsymbol{\theta}_y]\right\}p(b_{0i};\theta_b)db_{0i}\right]$$
$$+\sum_{i=1}^{n}(1-\omega_i)\log[p(v_i,\delta_i;\boldsymbol{\theta}_{z_0})],$$

is given by

$$l(\boldsymbol{\theta}) \approx \sum_{i=1}^{n}\omega_i\log\left[\sum_{k=1}^{Q}\frac{w_k}{\sqrt{\pi}}A_{ik}(\boldsymbol{\theta})B_{ik}(\boldsymbol{\theta})C_{ik}(\boldsymbol{\theta})\right] + \sum_{i=1}^{n}(1-\omega_i)\log[E_i(\boldsymbol{\theta})F_i(\boldsymbol{\theta})],$$

where

$$A_{ik}(\boldsymbol{\theta}) = \exp\left\{\delta_i\left[\log(\kappa_{ik1}(v_i)) - \frac{1}{2}\left(\log(8\pi) + \kappa_{ik2}^2(v_i)\right)\right]\right\},$$
$$B_{ik}(\boldsymbol{\theta}) = \exp\left\{(1-\delta_i)\log[1-\Phi(\kappa_{ik2}(v_i))]\right\},$$
$$C_{ik}(\boldsymbol{\theta}) = \exp\left\{\sum_{j=1}^{n_i}\log\left[\frac{1}{\sqrt{2\pi}\sigma_e}\exp\left\{-\frac{[y_i(t_{ij}) - (\boldsymbol{x}_{1i}(t_{ij})^\top\boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k)]^2}{2\sigma_e^2}\right\}\right]\right\},$$

with

$$\kappa_{ik1}(v_i) = \frac{2}{\alpha}\cosh\left(\frac{v_i - [\boldsymbol{x}_{2i}^\top\boldsymbol{\beta}_2 + \gamma(\boldsymbol{x}_{1i}(v_i)^\top\boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k)]}{2}\right),$$
$$\kappa_{ik2}(v_i) = \frac{2}{\alpha}\sinh\left(\frac{v_i - [\boldsymbol{x}_{2i}^\top\boldsymbol{\beta}_2 + \gamma(\boldsymbol{x}_{1i}(v_i)^\top\boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k)]}{2}\right),$$

with $s_k$ denoting the $k$-th root of the $Q$-th order Hermite polynomial and $w_k$ the corresponding weight,

$$E_i(\boldsymbol{\theta}) = \exp\left\{\delta_i\left[\log(\xi_{i1}(v_i)) - \frac{1}{2}\left(\log(8\pi) + \xi_{i2}^2(v_i)\right)\right]\right\},$$
$$F_i(\boldsymbol{\theta}) = \exp\left\{(1-\delta_i)\log[1-\Phi(\xi_{i2}(v_i))]\right\}.$$

# References

Albert, P. S. and Shih, J. H. (2010). On estimating the relationship between longitudinal measurements and time-to-event data using a simple two-stage procedure. *Biometrics 66*, 983-991.

Balakrishnan, N., Leiva, V., Sanhueza, A. and Vilca, F. (2009). Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM

algorithm. *Statistics and Operations Research Transactions 33*, 171-192.

Barros, M., Paula, G. A. and Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis 14*, 316-332.

Crowther, M. J., Abrams, K. R. and Lambert, P. C. (2012). Flexible parametric joint modelling of longitudinal and survival data. *Statistics in Medicine, 31*, 4456-4471

DeGruttola, V. and Tu, X. M. (1994). Modelling progression of CD-4 lymphocyte count and its relationship to survival time. *Biometrics 50*, 1003-1014.

Diggle, P., Farewell, D. and Henderson, R. (2007). Analysis of longitudinal data with drop-out: objectives, assumptions and a proposal. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 56*, 499-550.

Diggle, P. J., Sousa, I. and Chetwynd, A. G. (2008). Joint modelling of repeated measurements and time-to-event outcomes: the fourth Armitage lecture. *Statistics in Medicine 27*, 2981-2998.

Franco-Soto, D.C. (2014). PhD dissertation, Departamento de Estatística, Universidade de São Paulo (in Portuguese).
www.ime.usp.br/∼cpg/teses/Tese-DianaCarolinaFrancoSoto

Galea, M., Leiva, V. and Paula, G.A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal Applied Statistics 31*, 1049-1064.

Gay, D. (1990). *Usage Summary for Selected Optimization Routines*. AT&T Bell Laboratories, Murray Hill, NJ. Computing Science Technical Report No. 153.

Gould, A.L., Boye, M.E., Crowther, M.J., Ibrahim, J.G., Quartey, G., Micallef, S. and Bois, F.Y. (2014). Joint modelling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modelling working group. *Statistics in Medicine* doi:10.1002/sim.6141.

Greene, W. F. and Cai, J. (2004). Measurement error in covariates in the marginal hazards model for multivariate failure time data. *Biometrics 60*, 987-996.

Henderson, R., Diggle, P. and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics 1*, 465-480.

Hogan, J. W. and Laird, N. M. (1997a). Mixture models for the joint distributions of repeated measures and event times. *Statistics in Medicine 16*, 239-257.

Hogan, J. W. and Laird, N. M. (1997b). Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine 16*, 259-272.

Hsieh, F., Tseng, Y. K. and Wang, J. L. (2006). Joint modelling of survival and longitudinal data: likelihood approach revisited. *Biometrics 62*, 1037-1043.

Hu, P., Tsiatis, A. A. and Davidian, M. (1998). Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics 54*, 1407-1419.

Kalbfleisch, J.D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. 2nd edition. John Wiley & Sons, New Jersey.

Leiva, V., Barros, M.K., Paula, G.A. and Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis 51*, 5694-5707.

Leiva, V., Athayde, E., Azevedo, C. and Marchant, C. (2011). Modelling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter. *Journal of Applied Statistics 38*, 2819-2838.

Pawitan, Y. and Self, S. (1993). Modelling disease marker processes in AIDS. *Journal of the American Statistical Association 83*, 719-726.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. http://www.r-project.org/.

Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software 35*, 1-33.

Rizopoulos, D., Verbeke, G. and Lesaffre, E. (2009). Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society: Series B (Methodological) 71*, 637-654.

Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2010). Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival outcomes. *Biometrics 66*, 20-29.

Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics 67*, 819-829.

Rizopoulos, D. (2012a). *Joint Models for Longitudinal and Time-to-Event Data With Applications in R*. 1st edition. Chapman-Hall/CRC Biostatistics Series, Boca Raton.

Rizopoulos, D. (2012b). *JM: Shared Parameter Models for the Joint Modelling of*

*Longitudinal and Time-to-Event Data*. R package version 1.2-0. http://cran.r-project.org/web/packages/JM/index.html.

Schluchter, M. D. (1992). Methods for the analysis of informatively censored longitudinal data. *Statistics in Medicine 11*, 1861-1870.

Tsiatis, A.A. and Davidian, M. (2004). Joint modelling of longitudinal and time-to-event data: an overview. *Statistica Sinica 14*, 809-834.

Tsiatis, A. A., DeGruttola, V. and Wulfsohn, M. S. (1995). Modelling 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 90*, 27-37.

Wu, L. Liu, W., Yi, G. Y. and Huang, Y. (2012). Analysis of longitudinal and survival data: joint modelling, inference methods, and issues. *Journal of Probability and Statistics*, ID640153. doi:10.1155/2012/640153.

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics 53*, 330-339.

Ye, W., Lin, X. and Taylor J. M. G. (2008). Semiparametric modelling of longitudinal measurements and time-to-event data: a two-stage regression calibration approach. *Biometrics 64*, 1238-1246.

Yu, M., Law, N.J., Taylor, J.M.G. and Sandler, H.M. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica 14*, 835-862.