JMFit: A SAS Macro for Joint Models of Longitudinal and Survival Data

Danjie Zhang,
Gilead Sciences, Inc., 333 Lakeside Drive, Foster City, CA 94404, U.S.A., danjie.zhang@gilead.com

Ming-Hui Chen,
Department of Statistics, University of Connecticut, 215 Glenbrook Road U-4120, Storrs, CT 06269, U.S.A., ming-hui.chen@uconn.edu

Joseph G. Ibrahim,
Department of Biostatistics, University of North Carolina, McGavran Greenberg Hall CB#7420, Chapel Hill, NC 27599, U.S.A., ibrahim@bios.unc.edu

Mark E. Boye, and
Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, U.S.A., boyema@lilly.com

Wei Shen
Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, U.S.A., shen_wei_x1@lilly.com

Abstract

Joint models for longitudinal and survival data now have a long history of being used in clinical trials or other studies in which the goal is to assess a treatment effect while accounting for a longitudinal biomarker such as patient-reported outcomes or immune responses. Although software has been developed for fitting the joint model, no software packages are currently available for simultaneously fitting the joint model and assessing the fit of the longitudinal component and the survival component of the model separately as well as the contribution of the longitudinal data to the fit of the survival model. To fulfill this need, we develop a SAS macro, called JMFit. JMFit implements a variety of popular joint models and provides several model assessment measures including the decomposition of AIC and BIC as well as AIC and BIC recently developed in Zhang *et al.* (2014). Examples with real and simulated data are provided to illustrate the use of JMFit.

Keywords

AIC; BIC; Patient-reported outcome (PRO); Shared parameter model; Time-varying covariates

1. Introduction

The joint analysis of longitudinal and time-to-event outcomes has been widely published in statistical journals. One popular approach in joint modeling of longitudinal and survival data is based on shared random effects, where the longitudinal model and survival model share

common random effects and these random effects then induce correlation between the longitudinal and survival components of the model. This family of joint models is also called the "shared parameter models" (SPMs). There are two basic formulations of SPMs. The first is the "time trajectory model", denoted by SPM1, where one essentially substitutes the polynomial time trajectory function from the longitudinal model into the hazard function of the survival model, and in this case, the trajectory function acts like a time-varying covariate in the survival model. The second formulation, denoted by SPM2, is to directly include the random effects as covariates in the survival model. There are several R packages available in fitting joint models based on shared random effects, including **JM** (Rizopoulos 2012), **JMbayes** (Rizopoulos 2014), and **joineR** (Philipson *et al.* 2012). There is also a Stata module stjm (Crowther 2012; Crowther *et al.* 2013), which estimates shared random effects models. In addition, another R package, **lcmm** (Proust-Lima *et al.* 2014), estimates joint models based on shared latent classes.

One important issue in the joint modeling of longitudinal and survival data concerns the separate contribution of the model components to the overall goodness-of-fit of the joint model. Recently, Zhang *et al.* (2014) derived a novel decomposition of the AIC and BIC criteria into additive components that will allow us to assess the goodness of fit for each component of the joint model. Such a decomposition leads to the development of ΔAIC and ΔBIC, which quantify the change of AIC and BIC in fitting the survival data with and without using the longitudinal data. Thus, ΔAIC and ΔBIC can be used to determine the importance of the longitudinal data relative to the model fit of the survival data. In addition, ΔAIC and ΔBIC are also very useful in assessing whether a linear trajectory or quadratic trajectory is more suitable and also facilitating a direct comparison between SPM1's and SPM2's. These measures will help the data analyst in not only assessing each component of the joint model but also in determining the contribution of the longitudinal measures to the fit of the survival data. These newly developed model assessment criteria are not available in any of these packages or module mentioned before. We mention here that the methodology for ΔAIC and ΔBIC was fully developed in Zhang *et al.* (2014), but our goal here is the novel implementation of this methodology into user-friendly software along with a class of joint models for jointly analyzing longitudinal and time-to-event data.

This paper introduces JMFit, a SAS macro, that will allow us to fit the SPM1, SPM2, time-varying covariates, and two-stage models as well as to assess the goodness-of-fit of each of the longitudinal and survival components in the joint model. A detailed analysis of the longitudinal and survival data from a cancer clinical trial as well as an analysis of the simulated data are carried out to illustrate the functionality of JMFit. A detailed description of JMFit is given in Appendix A.

2. The models and model assessment

2.1. The joint models

Suppose that there are n subjects. For the ith subject, let $y_i(t)$ denote the longitudinal measure, which is observed at time $t \in \{a_{i1}, a_{i2}, \ldots, a_{im_i}\}$, where $0 \leq a_{i1} < a_{i2} < \cdots < a_{im_i}$ and $m_i \geq 1$. Here, $y_i(0)$ denotes the baseline value of the longitudinal measure. Let t_i and δ_i be the failure time and the censoring indicator such that $\delta_i = 1$ if t_i is a failure time and 0 if t_i

is right-censored for the ith subject. We further let $x_i(t)$ and z_i denote a p_L-dimensional vector of time-dependent covariates and a p_S-dimensional vector of baseline covariates, respectively. The joint model for (y_i, t_i) consists of the longitudinal component and the survival component.

For the longitudinal component, a mixed effects regression model is assumed for $y_i(t)$, which takes the form:

$$y_i(a_{ij}) = \boldsymbol{\theta}_i^\top \boldsymbol{g}(a_{ij}) + \boldsymbol{\gamma}^\top \boldsymbol{x}_i(t) + \epsilon_i(a_{ij}), \quad (1)$$

where $\boldsymbol{g}(a_{ij}) = \left(1, a_{ij}, a_{ij}^2, \ldots, a_{ij}^q\right)'$ is a polynomial vector of order q for $j = 1, \ldots, m_i$, $\boldsymbol{\theta}_i$ is a $(q+1)$-dimensional vector of random effects, and $\boldsymbol{\gamma}$ is a p-dimensional vector of regression coefficients. In (1), we further assume $\boldsymbol{\theta}_i \sim N(\boldsymbol{\theta}, \Omega)$, where $\boldsymbol{\theta}$ is the $(q+1)$-dimensional vector of the overall effects, Ω is a $(q+1) \times (q+1)$ positive definite covariance matrix with the lower triangle consisting of $\{\Omega_{00}, \Omega_{10}, \Omega_{11}, \ldots, \Omega_{qq}\}$, $\varepsilon_i(a_{ij}) \sim N(0, \sigma^2)$, and $\boldsymbol{\theta}_i$ and $\varepsilon_i(a_{ij})$ are independent. We note that in (1), if $q = 1$, $\boldsymbol{g}(a_{ij}) = (1, a_{ij})^\top$ and $\boldsymbol{\theta}_i^\top \boldsymbol{g}(a_{ij})$ represents a linear trajectory, and if $q = 2$, $\boldsymbol{g}(a_{ij}) = \left(1, a_{ij}, a_{ij}^2\right)^\top$ and $\boldsymbol{\theta}_i^\top \boldsymbol{g}(a_{ij})$ leads to a quadratic trajectory.

For the failure time t_i, we assume that the hazard function takes the form

$$\lambda(t|\lambda_0, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta}_i, \boldsymbol{g}(t), z_i) = \lambda_0(t)\ exp\left\{\beta\boldsymbol{\theta}_i^\top \boldsymbol{g}(t) + \boldsymbol{\alpha}^\top z_i\right\} \quad (2)$$

or

$$\lambda(t|\lambda_0, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta}_i, \boldsymbol{g}(t), z_i) = \lambda_0(t)\ exp\left\{\boldsymbol{\beta}^\top \boldsymbol{\theta}_i + \boldsymbol{\alpha}^\top z_i\right\} \quad (3)$$

where $\lambda_0(t)$ is the baseline hazard function, β is a one-dimensional regression coefficient in (2) and $\boldsymbol{\beta}$ is a $(q+1)$-dimensional vector of the regression coefficients in (3). Note that in (2) or (3), $\boldsymbol{\theta}_i$ and $\boldsymbol{g}(t)$ are the parameters and the functions from the longitudinal component of the joint model in (1) while λ_0, β (or $\boldsymbol{\beta}$), and $\boldsymbol{\alpha}$ are the only parameters pertaining to the survival component. As shown in Section 3.1, β (or $\boldsymbol{\beta}$) controls the association between the longitudinal marker and the time-to-event. A value of $\beta = 0$ (or $\boldsymbol{\beta} = \mathbf{0}$) implies no association between the longitudinal marker and the time-to-event. The joint model with hazard function specified in (2) is the trajectory model, denoted by SPM1, while the one with hazard function given by (3) is denoted by SPM2. Under SPM1, a positive value of β implies that a larger current value of the longitudinal marker is associated with a larger instantaneous hazard, whereas a negative value of β implies that a larger current value of the longitudinal marker is associated with a smaller instantaneous hazard.

2.2. The construction of the piecewise constant baseline hazard function

Assuming $\lambda_0(t)$ to be a piecewise constant baseline hazard function, we partition the time axis into J intervals with $0 = s_0^J < s_1^J < s_2^J < \ldots < s_{J-1}^J < s_J^J = \infty$. Then we assign a constant baseline hazard to each of the J intervals, that is,

$$\lambda_0(t) = \lambda_j, \quad t \in \left(s_{j-1}^J, s_j^J \right] \quad \text{for} \quad j = 1, 2, \ldots, J. \quad (4)$$

Let $t_1^* \leq t_2^* \leq \cdots \leq t_{n^*}^*$ be the n^* event times of the t_i's, where $n^* = \sum_{i=1}^{n} \delta_i$. We consider four algorithms to construct the s_j^J's.

Algorithm 1: Equally-Spaced Quantile Partition (ESQP)

Step 1: Compute $p_j = j/J$ for $j = 1, \ldots, J-1$.

Step 2: Let $n_j = [p_j n^*]$, which is the integer part of $p_j n^*$.

Step 3: Set

$$s_j^J = \begin{cases} \left(t_{n_j}^* + t_{n_{j+1}}^* \right)/2 & \text{if} \quad n_j = p_j n^*, \\ t_{n_{j+1}}^* & \text{if} \quad n_j < p_j n^*, \end{cases} \quad (5)$$

for $j = 1, \ldots, J-1$.

The ESQP is a popular approach to construct the piecewise constant hazard function, which is discussed in Ibrahim *et al.* (2001, Chapter 5) and also implemented in the R package **JM** developed by Rizopoulos (2010). We note that s_j^J is the p_jth quantile of the t_i^*'s and (5) is implemented in the **SAS UNIVARIATE** procedure as the default option for computing quantiles. We also note that the ESQP algorithm does not yield nested partitions in the sense that $\left\{ s_j^{J_1}, j = 1, \ldots, J_1 \right\}$ is not necessarily a subset of $\left\{ s_j^{J_2}, j = 1, \ldots, J_2 \right\}$ when $J_1 < J_2$. In order to construct partitions, we propose the following three bi-sectional quantile partition algorithms.

Algorithm 2: Left Bi-Sectional Quantile Partition (LBSQP)

Step 1: Decompose J into two parts:

$$J = 2^K + M, \quad (6)$$

where K and M are integers, and $M < 2^K$. In (6), $K = [\log J / \log 2]$, and then $M = J - 2^K$.

Step 2: Compute

(i) $a_k = k/2^K$, for $k = 1, \ldots, 2^K - 1$; and

      (ii) $b_m = (2m - 1)/2^{K+1}$, for $m = 1, \ldots, M(\quad 1)$.

Step 3: Sort $\{a_1, \ldots, a_{2^K - 1}, b_1, \ldots, b_M\}$ in ascending order and the resulting ordered $J - 1$ values are devoted by $p_1 \quad p_2 \quad \cdots \quad p_{J-1}$.

Step 4: Use Steps 2 and 3 of Algorithm 1 to compute $\left\{ s_j^J, j=1, \ldots, J - 1 \right\}$.

Algorithm 3: Middle Bi-Sectional Quantile Partition (MBSQP)

Step 1: The same as Algorithm 2.

Step 2: Compute

      (i) $a_k = k/2^K$, for $k = 1, \ldots, 2^K - 1$; and

      (ii)   for $m = 1, \ldots, M(\quad 1)$,

          (a) $b_m = (2^K - m)/2^{K+1}$, for $m = 1, 3, 5, \ldots$; and

          (b) $b_m = (2^K + m - 1)/2^{K+1}$, for $m = 2, 4, 6, \ldots$.

Steps 3 and 4: The same as Algorithm 2.

Algorithm 4: Right Bi-Sectional Quantile Partition (RBSQP)

Step 1: The same as Algorithm 2.

Step 2: Compute

      (i) $a_k = k/2^K$, for $k = 1, \ldots, 2^K - 1$; and

      (ii) $b_m = (2^{K+1} - (2m - 1))/2^{K+1}$, for $m = 1, \ldots, M(\quad 1)$.

Steps 3 and 4: The same as Algorithm 2.

If there are ties in $\left\{ s_j^J, j=1, \ldots, J \right\}$, say $s_\ell^J = s_{\ell+1}^J$, the interval $(s_\ell^J, s_{\ell+1}^J]$ is undefined. Thus we only use distinct values of the s_j^J. Let J_t denote the number of ties in $\left\{ s_j^J, j=1, \ldots, J \right\}$. Then the number of distinct intervals reduces to $J - J_t$.

Figure 1 shows how the partition intervals are constructed based on LBSQP, MBSQP, and RBSQP. Notice that when $J = 2^K$, $K = 1, 2, \ldots$, ESQP, LBSQP, MBSQP, and RBSQP yield the same partition. LBSQP, MBSQP, and RBSQP are desirable when there are more events at the beginning, in the middle, and at the end of the follow-up period, respectively. Another advantage of LBSQP, MBSQP, and RBSQP is that the resulting partitions are nested and, hence, the log-likelihood of the joint model increases in J when the longitudinal component remains fixed.

2.3. The joint likelihood

We rewrite (1) as follows:

$$\boldsymbol{y}_i = W_i \left(\boldsymbol{\theta}_i^\top, \boldsymbol{\gamma}^\top \right)^\top + \boldsymbol{\epsilon}_i,$$

where $\boldsymbol{y}_i = (y_i(a_{i1}), \ldots, y_i(a_{im_i}))^\top$, $W_i = \left(\left(g(a_{ij})^\top, \boldsymbol{x}_i(a_{ij})^\top \right)^\top, j = 1, \ldots, m_i \right)^\top$, and

$\boldsymbol{\epsilon}_i = (\epsilon_i(a_{i1}), \ldots, \epsilon_i(a_{im_i}))^\top \sim N\left(\mathbf{0}, \sigma^2 \mathbf{I_{m_i}} \right)$. The complete-data likelihood function of the longitudinal measures for the ith subject is given by

$$L\left(\boldsymbol{\gamma}, \sigma^2 | \boldsymbol{y_i}, \mathbf{W_i}, \boldsymbol{\theta_i}\right) = \frac{1}{(2\pi\sigma^2)^{\frac{m_i}{2}}} exp\left\{ -\frac{1}{2\sigma^2}\left(\boldsymbol{y}_i - W_i\left(\boldsymbol{\theta}_i^\top, \boldsymbol{\gamma}^\top\right)^\top \right)^\top \left(\boldsymbol{y}_i - W_i\left(\boldsymbol{\theta}_i^\top, \boldsymbol{\gamma}^\top\right)^\top \right) \right\},$$

(7)

for $i = 1, \ldots, n$. Note that the density of $\boldsymbol{\theta}_i$ is given by

$$f\left(\boldsymbol{\theta}_i | \boldsymbol{\theta}, \boldsymbol{\Omega}\right) = \frac{|\Omega|^{-\frac{1}{2}}}{(2\pi)^{\frac{q+1}{2}}} exp\left\{ -\frac{1}{2}(\boldsymbol{\theta}_i - \boldsymbol{\theta})^\top \Omega^{-1} (\boldsymbol{\theta}_i - \boldsymbol{\theta}) \right\}.$$

(8)

Let $\boldsymbol{\varphi} = (\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{a}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\theta}, \Omega)$. Using (2) (or (3)), (7), and (8), the observed-data likelihood function for $(\mathbf{y}_i, t_i, \delta_i)$ for the ith subject is given by

$$L\left(\boldsymbol{\varphi} | \boldsymbol{y_i}, \mathbf{t_i}, \delta_\mathbf{i}, z_\mathbf{i}, \mathbf{W_i}\right) = \int L\left(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha} | \mathbf{t_i}, \delta_\mathbf{i}, z_\mathbf{i}, \boldsymbol{\theta_i}, g\right) L\left(\boldsymbol{\gamma}, \sigma^2 | \boldsymbol{y_i}, \mathbf{W_i}, \boldsymbol{\theta_i}\right) f\left(\boldsymbol{\theta_i} | \boldsymbol{\theta}, \boldsymbol{\Omega}\right) d\boldsymbol{\theta}_i,$$

(9)

where the complete-data likelihood function for the survival component is written as

$$L\left(\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha} | \mathbf{t_i}, \delta_\mathbf{i}, z_\mathbf{i}, \boldsymbol{\theta_i}, g\right) = \left[\lambda\left(t_i | \lambda_0, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta_i}, g\left(\mathbf{t_i}\right), z_\mathbf{i}\right)\right]^{\delta_i} \times exp\left\{ -\int_0^{t_i} \lambda\left(u | \lambda_0, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta_i}, g\left(\mathbf{u}\right), z_\mathbf{i}\right) du \right\},$$

(10)

for $i = 1, \ldots, n$. In (2) (or (3)), when $\beta = 0$ (or $\boldsymbol{\beta} = \mathbf{0}$), the hazard function reduces to $\lambda(t | \lambda_0, \boldsymbol{a}, z_i) = \lambda_0(t) \exp(\boldsymbol{\alpha}^\top z_i)$. In this case, we fit the survival data alone and the likelihood function in (10) for the ith subject reduces to

$$L_0\left(\boldsymbol{\lambda}, \boldsymbol{\alpha} | \mathbf{t_i}, \delta_\mathbf{i}, z_\mathbf{i}\right) = \left\{\lambda_0(t_i) \, exp\left(\boldsymbol{\alpha}^\top z_i\right)\right\}^{\delta_i} exp\left[-exp\left(\boldsymbol{\alpha}^\top z_i\right) \left\{\int_0^{t_i} \lambda_0(u) \, du\right\}\right].$$

(11)

Letting $D_{\mathrm{obs}} = \{(\mathbf{y}_i, t_i, \delta_i, \boldsymbol{x}_i, z_i), i = 1, \ldots, n\}$ denote the observed data, the joint likelihood for all subjects is given by

$$L\left(\boldsymbol{\varphi}|\boldsymbol{g},\mathbf{D_{obs}}\right)=\prod_{i=1}^{n}L\left(\boldsymbol{\varphi}|\boldsymbol{y_i},\mathbf{t_i},\delta_\mathbf{i},\boldsymbol{z_i},\mathbf{W_i}\right).\tag{12}$$

2.4. AIC (BIC) decomposition and ΔAIC (ΔBIC)

Write $\boldsymbol{\varphi}_1 = (\boldsymbol{\gamma}, \sigma^2, \boldsymbol{\theta}, \Omega)$ and $\boldsymbol{\varphi}_2 = (\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\alpha})$. Let $f(\boldsymbol{\theta}_i|\boldsymbol{y}_i, W_i, \boldsymbol{\varphi}_1)$ be the conditional density of the random effects $\boldsymbol{\theta}_i$ given $\boldsymbol{y}_i$, and also let $L(\boldsymbol{\varphi}_1|\boldsymbol{y}_i, W_i) = \int L(\boldsymbol{\gamma}, \sigma^2|\boldsymbol{y}_i, W_i, \boldsymbol{\theta}_i) f(\boldsymbol{\theta}_i|\boldsymbol{\theta}, \Omega) d\boldsymbol{\theta}_i$, which is the likelihood function corresponding to the marginal distribution of $\boldsymbol{y}_i$. Following Zhang *et al.* (2014), the joint likelihood given in (12) can be decomposed as

$$L\left(\boldsymbol{\varphi}|\boldsymbol{g},\mathbf{D_{obs}}\right)=L_{Long}\left(\boldsymbol{\varphi}_1|\boldsymbol{g},D_{obs}\right)L_{Surv|Long}\left(\boldsymbol{\varphi}_2|\boldsymbol{g},\boldsymbol{\varphi}_1,D_{obs}\right),\tag{13}$$

where $L_{Long}\left(\boldsymbol{\varphi}_1|\boldsymbol{g},D_{obs}\right)=\prod_{i=1}^{n}L\left(\boldsymbol{\varphi}_1|\boldsymbol{y}_i,W_i\right)$ and
$L_{Surv|Long}\left(\boldsymbol{\varphi}_2|\boldsymbol{g},\boldsymbol{\varphi}_1,D_{obs}\right)=\prod_{i=1}^{n}\int L\left(\boldsymbol{\varphi}_2|t_i,\delta_i,\boldsymbol{z}_i,\boldsymbol{\theta}_i,\boldsymbol{g}\right)f\left(\boldsymbol{\theta}_i|\boldsymbol{y}_i,W_i,\boldsymbol{\varphi}_1\right)d\boldsymbol{\theta}_i$. Using (13), the decomposition of the total Akaike Information Criterion (AIC) (Akaike 1973) developed in Zhang *et al.* (2014) is given as

$$AIC=AIC_{Long}+AIC_{Surv|Long},$$

where $AIC= -2\,log\,L\,(\hat{\boldsymbol{\varphi}}|\boldsymbol{g},D_{obs})+2\,dim\,(\boldsymbol{\varphi})$,
$AIC_{Long}= -2\,log\,L_{Long}\,(\hat{\boldsymbol{\varphi}}_1|\boldsymbol{g},D_{obs})+2\,dim\,(\boldsymbol{\varphi}_1)$,
$AIC_{Surv|Long}= -2\,log\,L_{Surv|Long}\,(\hat{\boldsymbol{\varphi}}_2|\boldsymbol{g},\hat{\boldsymbol{\varphi}}_1,D_{obs})+2\,dim\,(\boldsymbol{\varphi}_2)$, and $\hat{\boldsymbol{\varphi}}$, $\hat{\boldsymbol{\varphi}}_1$, and $\hat{\boldsymbol{\varphi}}_2$ are the maximum likelihood estimates (MLEs) of $\boldsymbol{\varphi}$, $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$. Similarly, the total Bayesian Information Criterion (BIC) (Schwarz 1978) for the joint model can be decomposed into

$$BIC=BIC_{Long}+BIC_{Surv|Long},$$

where BIC = AIC+dim($\boldsymbol{\theta}$)(log n-2), BIC$_{Long}$ = AIC$_{Long}$+dim($\boldsymbol{\varphi}_1$)(log n–2), BIC$_{Surv|Long}$ = AIC$_{Surv|Long}$ + dim($\boldsymbol{\varphi}_2$)(log n–2). Using the decompositions of AIC and BIC, Zhang *et al.* (2014) proposed two new model assessment criteria given by

$$\Delta AIC= AIC_{Surv,0} - AIC_{Surv|Long},$$
$$\Delta BIC= BIC_{Surv,0} - BIC_{Surv|Long},$$

where

$$AIC_{Surv,0}= -2\sum_{i=1}^{n}log\,L_0\left(\hat{\boldsymbol{\lambda}},\hat{\boldsymbol{\alpha}}|t_i,\delta_i,\boldsymbol{z}_i\right)+2\,dim\,(\boldsymbol{\lambda},\boldsymbol{\alpha})\,,$$
$$BIC_{Surv,0}= -2\sum_{i=1}^{n}log\,L_0\left(\hat{\boldsymbol{\lambda}},\hat{\boldsymbol{\alpha}}|t_i,\delta_i,\boldsymbol{z}_i\right)+dim\,(\boldsymbol{\lambda},\boldsymbol{\alpha})\,log\,n,$$

and $L_0(\boldsymbol{\lambda}, \boldsymbol{a}|t_i, \boldsymbol{\delta}_i, z_i)$ is defined by (11). The ΔAIC or ΔBIC measure the gain of the fit in the survival component due to the longitudinal data with a penalty for the additional parameters in the survival component of the joint model. The model with a large value of ΔAIC (ΔBIC) is more preferred.

3. The SAS macro JMFit

3.1. Design

The SAS macro JMFit has been developed to assess model fit in joint models of longitudinal and survival data. In fact, it can fit five models, including the two types of joint models with linear and quadratic trajectories, as well as the time-varying covariates model. The macro JMFit consists of five submacros, SPM1L, SPM1Q, SPM2L, SPM2Q, and TVC, corresponding to the five models, respectively. The MODEL argument of JMFit specifies one of the following five models to be fitted:

SPM1L: SPM1 with Linear trajectory. The hazard function has the form

$$\lambda\left(t|\lambda_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}_\mathbf{i}, \boldsymbol{g}\left(\mathbf{t}\right), z_\mathbf{i}\right) = \lambda_0\left(t\right)\ exp\left\{\beta\left(\theta_{0i} + \theta_{1i}t\right) + \boldsymbol{\alpha}^\top z_i\right\}.$$

SPM1Q: SPM1 with Quadratic trajectory. The hazard function has the form

$$\lambda\left(t|\lambda_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}_\mathbf{i}, \boldsymbol{g}\left(\mathbf{t}\right), z_\mathbf{i}\right) = \lambda_0\left(t\right)\ exp\left\{\beta\left(\theta_{0i} + \theta_{1i}t + \theta_{2i}t^2\right) + \boldsymbol{\alpha}^\top z_i\right\}.$$

SPM2L: SPM2 with Linear trajectory. The hazard function has the form

$$\lambda\left(t|\lambda_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}_\mathbf{i}, \boldsymbol{g}\left(\mathbf{t}\right), z_\mathbf{i}\right) = \lambda_0\left(t\right)\ exp\left\{\beta_1\theta_{0i} + \beta_2\theta_{1i} + \boldsymbol{\alpha}^\top z_i\right\},$$

where θ_{0i} and θ_{1i} are subject-level random intercept and random slope.

SPM2Q: SPM2 with Quadratic trajectory. The hazard function has the form

$$\lambda\left(t|\lambda_0, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}_\mathbf{i}, \boldsymbol{g}\left(\mathbf{t}\right), z_\mathbf{i}\right) = \lambda_0\left(t\right)\ exp\left\{\beta_1\theta_{0i} + \beta_2\theta_{1i} + \beta_3\theta_{2i} + \boldsymbol{\alpha}^\top z_i\right\},$$

where θ_{0i}, θ_{1i}, and θ_{2i} are random effects (i.e., random intercept, random slope, and random quadratic coefficient).

TVC: Time-Varying Covariates model. The hazard function has the form

$$\lambda\left(t|\lambda_0, \beta, \boldsymbol{\alpha}, z_\mathbf{i}, \mathbf{y}_\mathbf{i}\left(\mathbf{t}\right)\right) = \lambda_0\left(t\right)\ exp\left\{\beta y_i\left(t\right) + \boldsymbol{\alpha}^\top z_i\right\},$$

where $y_i(t) = y_i(a_{ij})$ for $a_{ij} \leq t < a_{i,j+1}$ for $j = 1, \ldots, m_i$, where $a_{i,m_i+1} = \infty$. Note that the TVC model is a "non-joint" model, and the use of this model has great potential for bias (Fisher and Lin 1999).

We provide the two versions of SPM1L, SPM1Q, SPM2L, and SPM2Q with the TS argument. If TS is missing or equal to 0, the joint model will be fit; while "TS=1" yields the corresponding two-stage model. Similar to the method in Tsiatis *et al.* (1995), (i) we first fit the linear mixed model specified in (1) to the longitudinal data alone and then obtain the estimates of $\boldsymbol{\theta}_i$, denoted by $\hat{\boldsymbol{\theta}}_i$; and (ii) we replace $\boldsymbol{\theta}_i$ in (2) or (3) with the estimate $\hat{\boldsymbol{\theta}}_i$ at the second stage. The only difference between the joint model and the corresponding two-stage model is that $\boldsymbol{\theta}_i$ in (2) or (3) is replaced with the estimate $\hat{\boldsymbol{\theta}}_i$. This two-stage approach may potentially lead to biased and inefficient estimates (Ibrahim *et al.* 2010).

The number of intervals $J(\geq 1)$ for the piecewise constant baseline hazard function needs to be specified in the NPIECES argument. For the PARTITION argument, "1" represents "ESQP", "2" represents "LBSQP", "3" represents "MBSQP", and "4" represents "RBSQP".

JMFit automatically produces a rich text file (rtf) including five tables: (i) Number of Subjects; (ii) Fit Statistics; (iii) Survival Parameter Estimates (Survival Alone); (iv) Parameter Estimates; and (v) Hazard Ratios & λ Estimates.

3.2. Implementation details

If the observed longitudinal measures are sparse, the full trajectories of longitudinal measures might not be well estimated. For example, in the case of fitting a quadratic trajectory, the sign of the estimated second-order coefficient could be incorrect if the longitudinal measures were observed only within the first half of the follow-up period, leading to incorrect extrapolation when the observed progression time was far beyond the time of the last observed longitudinal measure.

Let $t_{\max,i} = \max_{1 \leq j \leq m_i}\{a_{ij}\}$. When $t > t_{\max,i}$, $y_i(t)$ is never observed and no longitudinal data are available to estimate the trajectory $\boldsymbol{\theta}_i^\top \boldsymbol{g}(t)$ for $t > t_{\max,i}$. Under SPM1, the extrapolation of the trajectory $\boldsymbol{\theta}_i^\top \boldsymbol{g}(t)$ beyond $t_{\max,i}$ may lead to a survival component of the joint model that fits the survival data poorly. In addition, such an extrapolation also causes a severe convergence problem in the SAS NLMIXED (SAS Institute Inc. 2011b) procedure especially when $t_{\max,i} << t_i$ for many subjects. To circumvent these issues, for SPM1, we modify the hazard function in (2) as

$$\lambda(t|\lambda_0, \boldsymbol{\alpha}, \beta, \boldsymbol{\theta}_i, \boldsymbol{g}, \boldsymbol{z}_i) = \lambda_0(t)\ exp\left\{\beta\boldsymbol{\theta}_i^\top \boldsymbol{g}\left(t - \left[t - t_{max,i}^*\right]_+\right) + \boldsymbol{\alpha}^\top \boldsymbol{z}_i\right\}, \quad (14)$$

or

$$\lambda\left(t|\lambda_0,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\theta_i},\boldsymbol{g},\boldsymbol{z_i}\right)=\lambda_0\left(t\right)\,exp\left\{\beta\boldsymbol{\theta}_i^\top\boldsymbol{g}\left(t-\left[t-t_{max,i}^*\right]_+\right)\times\frac{\tau-\left(t_{max,i}^*+\left[t-t_{max,i}^*\right]_+\right)}{\tau-t_{max,i}^*}+\boldsymbol{\alpha}^\top\boldsymbol{z}_i\right\},$$

(15)

where $\left[t-t_{max,i}^*\right]_+=max\left(t-t_{max,i}^*,0\right)$, $t_{max,i}^*=t_{max,i}+w\times max\left(t_i-t_{max,i},0\right)$, $w(\in$ [0,1]) is the proportion of $max(t_i-t_{max,i},0)$, and $\tau=\max_{1\ i\ n}\{t_i\}$, which is the last follow-up survival time. If $w=0$ (default), $t_{max,i}$ will be the starting point of the modified extrapolation of the trajectory; while $w=1$ implies that the trajectory extends to t_i with no $t_{max,i}$ adjustment. This modification can also be applied to the TS model corresponding to SPM1L or SPM1Q. There are two arguments called TMAXI and WEIGHT. If TMAXI is missing or equal to 0, no $t_{max,i}$ adjustment will be applied. If TMAXI=1, the $t_{max,i}$ adjustment based on hazard function given in (14) will be applied with weight given in the WEIGHT argument, that is, the trajectory will become flat after $t_{max,i}^*$. If TMAXI=2, the $t_{max,i}$ adjustment based on hazard function given in (15) will be applied with weight given in the WEIGHT argument, that is, starting at $t_{max,i}^*$, the trajectory will linearly go down to 0 at the last follow-up survival time. An "optimal" choice of weight w in conjunction with TMAXI option may be determined by either $AIC_{Surv|Long}$ or $BIC_{Surv|Long}$. The purpose of defining $t_{max,i}$ is that we create an extrapolation of the longitudinal measures so that the trajectory function can be well estimated. This extrapolation is needed when there are few longitudinal measures at later points. We note here that the $t_{max,i}$ method corresponds to a "Prediction Carried Forward" approach, which may potentially induce bias in the estimation. We also note that this is not an issue for the SPM2, in which the hazard function is independent of $\boldsymbol{g}(t)$.

The OPTIONS argument allows users to specify options (e.g., integration method, optimization technique, and convergence criteria) that are available in the PROC NLMIXED statement. If OPTIONS is missing, JMFit will use adaptive Gaussian quadrature to approximate the integral of the likelihood over the random effects, perform a quasi-Newton optimization, and apply a relative gradient convergence criterion of 10^{-8}. All the methods mentioned above are default methods in PROC NLMIXED. The Riemann integral is used to compute the cumulative hazard function for the trajectory models. Each time interval is divided into 200 subintervals.

A big challenge in fitting joint models using the SAS NLMIXED procedure is convergence. Poor initial values may lead to the failure of convergence in NLMIXED. To address this issue, we first fit the longitudinal data alone using the SAS MIXED procedure to obtain the estimates, $\hat{\boldsymbol{\gamma}},\hat{\sigma}^2,\hat{\boldsymbol{\theta}},\hat{\Omega},\hat{\boldsymbol{\theta}}_i$, for the parameters in the longitudinal component of the joint model. Using $(\hat{\boldsymbol{\theta}}_i,i=1,\ldots,n)$ to replace $(\boldsymbol{\theta}_i,i=1,\ldots,n)$ in (2) (or (3)), we fit the survival data alone to obtain the estimates, $\hat{\beta},$ (or $\hat{\boldsymbol{\beta}}),\hat{\boldsymbol{\alpha}}$, and $\hat{\boldsymbol{\lambda}}$, for the parameters in the survival

component of the joint model. Finally, these estimates are used as the initial values for the joint model.

JMFit does not exclude any longitudinal measures for the joint models. If one wishes to exclude the longitudinal measures observed after the survival time t_i, those longitudinal measures should be pre-excluded in the input longitudinal data for JMFit. "CAUTION: Longitudinal measures are observed after the survival time." will be given at the end of the output file if there are any longitudinal measures observed after the survival time.

4. Examples

4.1. The EMPHACIS data

To illustrate how JMFit works, we use the data from a multicenter, randomized, single-blind, **EMPHACIS** lung cancer clinical trial (**E**valuation of **M**TA in Mesothelioma in a **P**hase III Study with **Cis**platin). The study drug was a **m**ulti-**t**argeted **a**ntifolate (MTA) pemetrexed which was given in combination with cisplatin (the PEM/Cis arm), and the active-treatment comparator was cisplatin alone (the Cis arm). The treatment for both arms was structured as six 21-day cycles of therapy; patients receiving treatment benefit could receive additional cycles based on investigator discretion.

Malignant pleural mesothelioma is a rapidly progressing and highly symptomatic malignancy with a median survival time of 6 to 9 months. Accordingly, patient-reported assessments are important for evaluation of disease progression and patients' response to therapy. In oncology, the patients' importance ratings on the magnitude of progression-free survival improvement have been shown to depend on the severity of disease-related symptoms (Bridges *et al.* 2012). We analyzed the disease-specific patient-reported Lung Cancer Symptom Scales (LCSS) to evaluate the patient-level association of five instrument items, anorexia, cough, dyspnea, fatigue, and pain, with progression-free survival using EMPHACIS trial data.

We consider the same subset of the EMPHACIS data as in Zhang *et al.* (2014), which consists of 425 patients with at least one post-baseline value of each longitudinal measure and seven binary covariates, including therapy (1=pemetrexed/cisplatin and 0=cisplatin alone), race (1=white and 0=others), gender (1=male and 0=female), age (1=' 65' and 0='< 65'), Karnofsky (1='90-100' and 0='<90'), stage (1=stage I/II and 0=stage III/IV), and bf (1=full vitamin supplementation and 0=other). The detailed description of this study can be found in Zhang *et al.* (2014). By applying joint models in this study, we compare the longitudinal LCSS symptoms in terms of their contribution to the overall fit of survival data via AIC and BIC, which are computed using JMFit.

Suppose we fit the trajectory model with a linear trajectory and $J = 3$ for pain. Then we create two data sets named as "lcss_long_pain" and "lcss_surv_pain" for the macro JMFit's LONG and SURV options, respectively. The MODEL option is set to "SPM1L" and TS is set to 0. We assign 1 to TMAXI for $t_{\max,i}$ adjustment with WEIGHT=0. The NPIECES option is set to 3 and the PARTITION option is set to 2 for the LBSQP algorithm.

JMFit is called:

```
%JMFit(LONG = lcss_long_pain, SURV = lcss_surv_pain, MODEL = SPM1L, TS = 0,
TMAXI = 1, WEIGHT = 0, NPIECES = 3, PARTITION = 2);
```

The rtf file "Output for pain under SPM1L with $J = 3$ (Partition=2).rtf" is generated by JMFit, which lists five tables. Table "Number of Subjects" given in Figure 2 shows that there are 425 patients in data sets lcss_long_pain and lcss_surv_pain, respectively, and 425 subjects are used, implying that the id's of the subjects in these two data sets match.

Figure 3 shows the "Fit Statistics" table, which consists of the log likelihood, AIC_{Long} (BIC_{Long}), AIC_{Surv} (BIC_{Surv}), and AIC (BIC). Inside the JMFit macro, PROC NLMIXED provides the log likelihood, AIC and BIC, and PROC IML (SAS Institute Inc. 2011a) is used to compute AIC_{Long} and BIC_{Long}.

Next, the table, titled "Survival Parameter Estimates (Survival Alone)" shown in Figure 4, is obtained by fitting the survival data alone. For this table, the estimate, the standard error (SE), the degrees of freedom (DF), the t-value, the p-value, and the 95% confidence interval (CI) are all shown for each parameter. In addition, the gradient of the negative log-likelihood function is displayed, which can be used to check convergence of PROC NLMIXED. A small gradient implies better convergence. The largest absolute value of the gradients shown in Figure 4 is 0.00322, indicating good convergence of PROC NLMIXED.

PROC NLMIXED produces the table titled "Parameter Estimates" in Figure 5, which consists of three subtables: "Covariance Parameter Estimates", "Longitudinal Parameter Estimates", and "Survival Parameter Estimates". For each parameter in this table, the estimate, SE, DF, t value, p value, 95% CI, and gradient are all provided. The Subtable "Covariance Parameter Estimates" consists of the estimates of the three lower-triangle elements (Ω_{00}, Ω_{10}, Ω_{11}) of the random-effects covariance matrix as well as the estimate of the standard deviation (σ) for the error term. The estimates of the coefficients of the variables from the longitudinal component are shown in Subtable "Longitudinal Parameter Estimates". In Subtable "Survival Parameter Estimates", "Param/Var" lists all the names of the variables as well as the parameters from the survival component. In this subtable, $\log \lambda_1$, $\log \lambda_2$, and $\log \lambda_3$ are the natural logarithms of the piecewise baseline hazards for the three intervals, respectively. The hazard ratios of the covariates and β as well as the estimates of λ are shown in Figure 6. From Figure 5, we see that therapy has p-values of 0.1451 and <.0001 in the longitudinal and survival submodels, respectively, indicating that therapy is not statistically significant at the 0.05 level in the longitudinal submodel and it is highly significant in the survival submodel. Also, from Figure 6, the estimated hazard ratio of therapy in the survival model is 0.6272 with a 95% CI of (0.508, 0.775), implying that the treatment reduces the risk of disease progression by about 37%. From Figure 5, we also see that (i) karnofsky is highly significant in both the longitudinal and survival submodels and (ii) the other significant variables in the survival submodel include the coefficient β corresponding to the linear trajectory and stage. A significant coefficient β indicates that pain is highly associated with disease progression.

From the "Fit Statistics" tables in Figure 3 and Figure 7, we see that pain had the largest values of AIC and BIC and cough had the smallest values of AIC and BIC, which indicate that pain led to the most gain in fitting the PFS data while cough had the least contribution to the fit of the PFS data.

Since $AIC = AIC_{Surv,0} - AIC_{Surv|Long}$, both $AIC_{Surv,0}$ and $AIC_{Surv|Long}$ change when J changes. Thus, the "best J" based on AIC may not necessarily lead to the best survival submodel or the best joint model. As an illustration, Table 1 shows the values of $AIC_{Surv|Long}$ and AIC for pain under SPM1L with different partition algorithms and different Js. From Table 1, we see that (i) $J = 7$ (or $J = 6$) is the "best" choice according to AIC; (ii) $J = 2$ is the "best" one based on $AIC_{Surv|Long}$; and (iii) the values of $AIC_{Surv|Long}$ are 2175.29, 2175.29, and 2175.60 for LBSQP, MBSQP, and RBSQP when $J = 7$, which are much larger than those when $J = 2$. Since our goal is to fit the joint model to both the longitudinal data and the survival data, it is more appropriate to use $AIC_{Surv|Long}$ to determine the number of intervals for the piecewise constant baseline hazard function.

Table 2 shows the values of $AIC_{Surv|Long}$ for cough and pain under SPM1L with $t_{max,i}$ adjustments using different w's. We can see from Table 2 that, among the five values of w, for both $t_{max,i}$ adjustments, cough and pain have the smallest values of $AIC_{Surv|Long}$ at $w = 0.25$ and 0, respectively. For cough, the two $t_{max,i}$ adjustments result in similar values of $AIC_{Surv|Long}$; and for pain, TMAXI=2 slightly outperforms TMAXI=1. In addition, both cough and pain have the largest values of $AIC_{Surv|Long}$ at $w = 1$, indicating that the full trajectories were not well estimated.

4.2. A simulated data example

Since we do not have permission to distribute the EMPHACIS dataset used in Section 4.1 publicly, we consider a simulated data example, in which the simulated data can be downloaded directly from the journal's website.

We generated a simulated data set with $n = 400$ subjects as follows. First, the time points (a_{ij}'s) at which the longitudinal measures were taken were fixed at (0, 21, 42, 63, 84, 105, 126)/30.4375. For each subject, we generated seven binary covariates, ($x_{i1}, \ldots, x_{i7}$), independently from Bernoulli distributions with success probabilities (i.e., $P(x_{ij} = 1)$, $j = 1, \ldots, 7$)) 0.49, 0.92, 0.81, 0.49, 0.38, 0.56, and 0.74, respectively. These proportions were estimated from the EMPHACIS data, corresponding to the covariates treatment, race/ethnicity, gender, age, Karnofsky status, baseline stage of disease, and vitamin supplementation, respectively. Second, we simulated the longitudinal trajectory as

$$\mu_i(a_{ij}) = (\theta_0 + \theta_{0i}) + (\theta_1 + \theta_{1i})\, a_{ij} + \gamma x_{\mathbf{i}},$$

where $\theta_0 = 0.62$, $\theta_1 = 0.04$, $\boldsymbol{\gamma} = (-\mathbf{0.11}, -\mathbf{0.10}, -\mathbf{0.18}, -\mathbf{0.06}, -\mathbf{0.58}, -\mathbf{0.09}, \mathbf{0.10})^\top$, and

$(\theta_{0i}, \theta_{1i})^\top \sim N\left(0, \begin{pmatrix} 0.62 & -0.04 \\ -0.04 & 0.06 \end{pmatrix}\right)$. Finally, we generated the longitudinal data from a $N(\mu_i(a_{ij}), \sigma^2)$ distribution with $\sigma = 0.54$ and t_i^* as

$$t_i^* = -\frac{log\,(1-U)}{\lambda\,exp\,\{\beta_1\,(\theta_0+\theta_{0i})+\beta_2\,(\theta_1+\theta_{1i})+\boldsymbol{\alpha}\boldsymbol{x_i}\}},$$

where $\boldsymbol{\alpha}=(-0.36, 0.15, 0.04, -0.003, -0.33, -0.38, -0.07)^\top$, $\beta_1 = 0.26$, $\beta_2 = 1.17$, $\lambda = $ exp(-1.67), and $U \sim U(0,1)$. This longitudinal dataset is denoted by D_{Long}. Note that the values of the parameters were obtained by fitting SPM2L to the EMPHACIS data in which the longitudinal measure corresponds to pain. The censoring time C_i was generated from an exponential distribution with mean 50, and the right-censoring percentage was about 12%. The failure time and censoring indicator were calculated as $t_i = min\,\{t_i^*, C_i\}$ and $\delta_i = 1\,\{t_i^* \leq C_i\}$, where $1\{A\}$ denotes the indicator function such that $1\{A\} = 1$ if A is true and 0 otherwise. This survival dataset is denoted by D_{Surv}.

We also generated three additional sets of longitudinal data, with longitudinal trajectories simulated from

$$\mu_{\ell i}\,(a_{ij}) = (\theta_0+\theta_{0i}+\tau_{\ell 0i}) + (\theta_1+\theta_{1i}+\tau_{\ell 1i})\,a_{ij}+\boldsymbol{\gamma}\boldsymbol{x_i},$$

where (i) $(\tau_{10i}, \tau_{11i})^\top \sim N\left(\mathbf{0}, \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{pmatrix}\right)$; (ii) $(\tau_{20i}, \tau_{21i})^\top \sim N\left(\mathbf{0}, \begin{pmatrix} 0.5^2 & 0 \\ 0 & 0.5^2 \end{pmatrix}\right)$; and (iii) $(\tau_{30i}, \tau_{31i})^\top \sim N\left(\mathbf{0}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$. Then the longitudinal data were generated from a $N(\mu_{\ell i}(a_{ij}), \sigma^2)$ distribution with $\sigma = 0.54$ for $\ell = $ 1, 2, 3. These three additional sets of longitudinal data were coupled with the same survival times as in $D_{Long} + D_{Surv}$ to form three additional data sets. These resulting data sets are denoted by $D_{Long1} + D_{Surv}$, $D_{Long2} + D_{Surv}$, and $D_{Long3} + D_{Surv}$. This simulation setting is similar to Simulation III in Zhang *et al.* (2014) except for the six additional covariates.

Figure 8 shows the fit statistics for $D_{Long} + D_{Surv}$, $D_{Long1} + D_{Surv}$, $D_{Long2} + D_{Surv}$, and $D_{Long3} + D_{Surv}$, respectively, using JMFit. We see from Figure 8 that $D_{Long} + D_{Surv}$ has the largest values of ΔAIC and ΔBIC.

The rest of the output for $D_{Long} + D_{Surv}$ are provided in Appendix B. The output for $D_{Long\ell}+D_{Surv}$ ($\ell = 1, 2, 3$) are omitted here for brevity.

5. Concluding remarks

The JMFit SAS macro fits the joint models for longitudinal and survival data. The piecewise exponential constant hazard model is assumed for the baseline hazard function. The time-axis is partitioned into J intervals, which are constructed by four algorithms, namely, ESQP, LBSQP, MBSQP, and RBSQP. JMFit allows users to specify the number of intervals J and the partition method. Five models, including SPM1L, SPM1Q, SPM2L, SPM2Q, and TVC, are implemented in JMFit. This SAS macro also computes AIC, BIC, ΔAIC, ΔBIC, and the estimates of the parameters in the joint model. The computational time of JMFit depends on which of SPM1L, SPM1Q, SPM2L, SPM2Q, or TVC is chosen and how big the dataset is.

For the example given in Section 4.1, it took 42 seconds to fit SPM1L with $J = 3$ based on LBSQP for pain on a Dell PC with an Intel i5 processor, 3.30 GHz CPU, and 8 GB of memory. On the same PC, it only took 10 to 11 seconds to fit SPM2L to each of the simulated datasets illustrated in Section 4.2.

The JMFit SAS macro provides two versions for SPM1L, SPM1Q, SPM2L, and SPM2Q with the TS argument, with "TS=1" yielding the corresponding two-stage model. Comparing these models with TS missing or equal to 0, for the two-stage models, (i) we first fit (1) to the longitudinal data alone and obtain the estimates of $\boldsymbol{\theta}_i$, denoted by $\hat{\boldsymbol{\theta}}_i$; and (ii) we then use $\hat{\boldsymbol{\theta}}_i$ in (2) and (3). The TS models are also substantially different than the TVC model since the TS models do not require the LOCF assumption.

The current version of the JMFit SAS macro only fits linear and quadratic models for the longitudinal outcome and the piecewise constant baseline hazard function for the survival submodel. In the joint modeling framework, other dependence structures, such as dependence through the derivatives of the trajectory function or interactions with covariates as well as spline approximations to the baseline hazard could be assumed. In addition, other trajectory functions may be more appropriate to model the time effect on the longitudinal outcomes in certain applications. These additional features could be built in the JMFit macro in a future release.

Finally, we note that the EMPHACIS dataset used in this paper is proprietary and cannot be distributed publicly. However, the simulated datasets D_{Surv}, D_{Long}, D_{Long1}, D_{Long2}, and D_{Long3} are available for downloading from the journal website.

Acknowledgments

A. The macro JMFit

The SAS macro JMFit as well as the five submacros should be stored in a folder named "jmfit". Then JMFit can be accessed by including the following lines:

```
filename jmfit "directory of the file JMFit";
%include jmfit(JMFit.sas);
%JMFit(LONG=, SURV=, MODEL=, TS=, TMAXI=, WEIGHT=, NPIECES=, PARTITION=,
OPTIONS=, INITIAL=, OUTPUT=);
```

Inputs for JMFit

LONG: Data set with the first three columns, SID (subject ID), Y (longitudinal measure), A (time at which Y was taken), and additional columns for covariates (X_{L1}-X_{Lp}), where SID, Y, and A should be arranged in the first, second, and third columns, and $X_{L1}, \cdots, X_{Lp}$ should

be placed after column 3, which can be enforced in SAS by using the "retain" command. Note that $X_{L1}, ..., X_{Lp}$ can be time-dependent or baseline covariates. Required.

SURV: Data set with the first three columns, SID, survival time (T), censoring indicator (delta) (1 = death and 0 = censored), and additional columns for covariates (X_{S1}-X_{Sq}), where SID, T, and delta should be arranged in the first, second, and third columns, and $X_{S1}, \ldots, X_{Sq}$ should be placed after column 3. Required.

MODEL: Model specification. Required. One of

> **1.** SPM1L: Shared Parameter Model 1 with Linear trajectory.
>
> **2.** SPM1Q: Shared Parameter Model 1 with Quadratic trajectory.
>
> **3.** SPM2L: Shared Parameter Model 2 with Linear trajectory.
>
> **4.** SPM2Q: Shared Parameter Model 2 with Quadratic trajectory.
>
> **5.** TVC: Time-Varying Covariates Model.

TS: Indicates whether to implement the model specified in the MODEL argument or the corresponding two-stage model. If 0 (default), the model specified in the MODEL argument will be fit. If 1, the corresponding two-stage model will be fit instead. It only works for SPM1L, SPM1Q, SPM2L, and SPM2Q.

TMAXI: $t_{\max,i}$ adjustment to the model specified in the MODEL argument. If 0 (default), no $t_{\max,i}$ adjustment will be applied. If 1, the trajectory will become flat after $t^*_{max,i} = t_{max,i} + WEIGHT \times max(t_i - t_{max,i}, 0)$. If 2, starting at $t^*_{max,i}$, the trajectory will linearly go down to 0 at the last follow-up survival time. It only works for SPM1L and SPM1Q.

WEIGHT: The proportion ($\in [0, 1]$) of $max(t_i - t_{\max,i}, 0)$. If 0 (default), the starting point of the modified extrapolation of the trajectory is $t_{\max,i}$. If 1, the trajectory extends to t_i with no $t_{\max,i}$ adjustment. It only works when TMAXI=1 or TMAXI=2.

NPIECES: Number of intervals $J(\ 1)$ for the piecewise constant baseline hazard function.

Required.

PARTITION: Algorithm for constructing the partition of the time axis. Required. One of

> **(i)** 1: Equally-Spaced Quantile Partition (ESQP).
>
> **(ii)** 2: Left Bi-Sectional Quantile Partition (LBSQP).
>
> **(iii)** 3: Middle Bi-Sectional Quantile Partition (MBSQP).
>
> **(iv)** 4: Right Bi-Sectional Quantile Partition (RBSQP).

OPTIONS: Allows users to specify options that are available in the PROC NLMIXED statement. For example, "OPTIONS=%str(QPOINTS=5 TECH=CONGRA ABSGCONV=0.0001)" specifies Gaussian quadrature with five quadrature points for

approximating the integral of the likelihood over the random effects, the conjugate-gradient optimization, and an absolute gradient convergence criterion of 0.0001.

INITIAL: Allows users to set their own initial values. JMFit will automatically generate the starting values for the model parameters and these initial values will be stored in the data set "_initial". Since the order of the parameters is very important when calculating AIC_{Long} and BIC_{Long}, users are recommended to change the initial values in "_initial" and then rename "_initial".

OUTPUT: Name of the output rich text file (rtf). One can also specify the directory in which the file will be put. For example, the output file named "myoutput" will be stored in "C: \ . . . " by "%JMFit(..., OUTPUT = C: \. . . \myoutput);" If OUTPUT is not specified, the file will be indexed by the name of Y from LONG and the model's name.

> Note 1: (i) The name of the SID variable in LONG should be the same as that of the SID variable in SURV; (ii) A and T should be in the same unit of time (month preferred); (iii) the categorical covariates must be coded as dummy variables; and (iv) the SAS macro allows for any numbers of covariates for both components of the joint model and the covariates for the longitudinal component can be totally different from those for the survival component.

> Note 2: (i) "ERROR: Not enough memory to generate code." This might occur if J is too big; (ii) a too long path for the OUTPUT may lead to an error; (iii) the macro is assuming "options validvarname=v7;" for valid variable names that can be processed in SAS; and (iv) the calculations of AIC_{Long} and BIC_{Long} require the IML Procedure.

> Note 3: No missing values are allowed in both data sets.

Output for JMFit

The macro automatically produces an rtf file indexed by the name of Y from LONG and the model's name. The rtf file includes five tables: (i) Number of Subjects; (ii) Fit Statistics; (iii) Survival Parameter Estimates (Survival Alone); (iv) Parameter Estimates; and (v) Hazard Ratios & λ Estimates.

Note: The construction of the Parameter Estimates table is different for each model. For SPM1L, SPM1Q, SPM2L, and SPM2Q, it consists of three subtables: "Covariance Parameter Estimates", "Longitudinal Parameter Estimates", and "Survival Parameter Estimates"; for the TS model corresponding to SPM1L, SPM1Q, SPM2L, or SPM2Q, there are two tables: "Longitudinal Parameter Estimates (Stage I)" and "Survival Parameter Estimates (Stage II)"; and for the TVC model, there is only one table named "Parameter Estimates".

B. Output for the simulated data DLong + DSurv under SPM2L

Number of Subjects	
Subjects_in_long	400
Subjects_in_surv	400
Subjects_Used	400

Figure 9.
The number of subjects for $D_{\text{Long}} + D_{\text{Surv}}$.

Survival Parameter Estimates (Survival Alone)							
Param/Var	Estimate	SE	DF	T-Value	P-Value	95% CI	Gradient
therapy	-0.3373	0.1081	400	-3.12	0.0019	(-0.550, -0.125)	0.000045
race	-0.00216	0.1811	400	-0.01	0.9905	(-0.358, 0.354)	0.000217
gender	0.05369	0.1272	400	0.42	0.6732	(-0.196, 0.304)	0.001135
age	0.04122	0.1078	400	0.38	0.7023	(-0.171, 0.253)	-0.00013
karnofsky	-0.2666	0.1116	400	-2.39	0.0173	(-0.486, -0.047)	-0.00031
stage	-0.2204	0.1088	400	-2.03	0.0435	(-0.434, -0.006)	-0.00028
bf	-0.03884	0.1262	400	-0.31	0.7585	(-0.287, 0.209)	-0.0001
$\log \lambda_1$	-1.5882	0.2502	400	-6.35	<.0001	(-2.080, -1.096)	0.000048

Figure 10.
The estimates of the parameters obtained by fitting the survival data alone for $D_{\text{Long}} + D_{\text{Surv}}$.

Hazard Ratios & λ Estimates		
Parameter	**Estimate**	**95% CI**
HR_therapy	0.7020	(0.563, 0.875)
HR_race	1.0552	(0.732, 1.521)
HR_gender	0.9778	(0.756, 1.264)
HR_age	1.1057	(0.889, 1.375)
HR_karnofsky	0.7499	(0.598, 0.940)
HR_stage	0.8237	(0.661, 1.026)
HR_bf	0.9328	(0.722, 1.205)
λ_1	0.1809	(0.109, 0.301)
HR_β_1	1.2041	(1.046, 1.387)
HR_β_2	3.6966	(2.312, 5.910)

Figure 11.
The hazard ratios and λ estimates for $D_{\text{Long}} + D_{\text{Surv}}$ under SPM2L.

Param/Var	Estimate	SE	DF	T-Value	P-Value	95% CI	Gradient
Parameter Estimates							
Covariance Parameter Estimates							
Ω_{00}	0.6819	0.05753	398	11.85	<.0001	(0.569, 0.795)	0.005593
Ω_{10}	-0.05231	0.01440	398	-3.63	0.0003	(-0.081, -0.024)	0.022189
Ω_{11}	0.06822	0.006320	398	10.79	<.0001	(0.056, 0.081)	-0.00706
σ	0.5252	0.008301	398	63.26	<.0001	(0.509, 0.542)	0.000563
Longitudinal Parameter Estimates							
Intercept	0.6852	0.1916	398	3.58	0.0004	(0.309, 1.062)	-0.00962
longt	0.03229	0.01491	398	2.17	0.0309	(0.003, 0.062)	-0.03217
therapy	-0.01673	0.08539	398	-0.20	0.8447	(-0.185, 0.151)	-0.00482
race	-0.1503	0.1449	398	-1.04	0.3002	(-0.435, 0.135)	-0.00991
gender	-0.1901	0.1019	398	-1.87	0.0628	(-0.390, 0.010)	-0.00488
age	-0.1539	0.08503	398	-1.81	0.0711	(-0.321, 0.013)	-0.02083
karnofsky	-0.5958	0.08773	398	-6.79	<.0001	(-0.768, -0.423)	-0.00258
stage	-0.2010	0.08602	398	-2.34	0.0199	(-0.370, -0.032)	0.011007
bf	-0.07759	0.09840	398	-0.79	0.4309	(-0.271, 0.116)	-0.00643
Survival Parameter Estimates							
therapy	-0.3539	0.1118	398	-3.16	0.0017	(-0.574, -0.134)	-0.01544
race	0.05376	0.1859	398	0.29	0.7726	(-0.312, 0.419)	0.000899
gender	-0.02245	0.1307	398	-0.17	0.8637	(-0.279, 0.234)	0.008303
age	0.1005	0.1109	398	0.91	0.3656	(-0.118, 0.319)	-0.0022
karnofsky	-0.2878	0.1150	398	-2.50	0.0128	(-0.514, -0.062)	0.010682
stage	-0.1940	0.1117	398	-1.74	0.0831	(-0.414, 0.026)	-0.00048
bf	-0.06954	0.1302	398	-0.53	0.5937	(-0.326, 0.186)	0.001669
$\log \lambda_1$	-1.7098	0.2590	398	-6.60	<.0001	(-2.219, -1.201)	-0.00608
β_1	0.1858	0.07175	398	2.59	0.0100	(0.045, 0.327)	0.003305
β_2	1.3074	0.2387	398	5.48	<.0001	(0.838, 1.777)	0.00382

Figure 12.
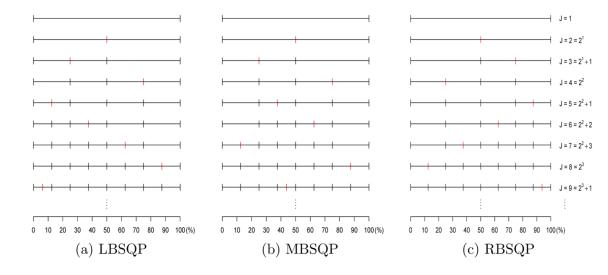The estimates of the parameters for $D_{\text{Long}} + D_{\text{Surv}}$ under SPM2L.

References

Akaike, H. Information Theory and an Extension of the Maximum Likelihood Principle.. In: Petrov, BN.; Csaki, F., editors. Proceedings of the Second International Symposium on Information Theory. Akademiai Kiado; Budapest: 1973. p. 267-281.

Bridges JFP, Mohamed AF, Finnern HW, Woehl A, Hauber AB. Patients' Preferences for Treatment Outcomes for Advanced Non-Small Cell Lung Cancer: A Conjoint Analysis. Lung Cancer. 2012; 77:224–231. [PubMed: 22369719]

Crowther, MJ. STJM: Stata Module to Fit Shared Parameter Joint Models of Longitudinal and Survival Data. 2012. URL http://econpapers.repec.org/software/bocbocode/s457502.htm/

Crowther MJ, Abrams KR, Lambert P. Joint Modeling of Longitudinal and Survival Data. The Stata Journal. 2013; 13:165–184.

Fisher LD, Lin D. Time-Dependent Covariates in the Cox Proportional-Hazards Regression Model. Annual Review of Public Health. 1999; 20:145–157.

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

Ibrahim JG, Chu H, Chen LM. Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. Journal of Clinical Oncology. 2010; 28:2796–2801. [PubMed: 20439643]

Philipson, P.; Sousa, I.; Diggle, P.; Williamson, P.; Kolamunnage-Dona, R.; Henderson, R. joineR: Joint Modelling of Repeated Measurements and Time-to-Event Data.. R package version 1.0-3. 2012. URL http://cran.r-project.org/web/packages/joineR/index.html.

SAS Institute Inc. SAS/IML Software, Version 9.3. Cary, NC: 2011a. URL http://www.sas.com/

SAS Institute Inc. SAS/STAT Software, Version 9.3. Cary, NC: 2011b. URL http://www.sas.com/

Proust-Lima, C.; Philipps, V.; Diakite, A.; Liquet, B. lcmm: Estimation of Extended Mixed Models Using Latent Classes and Latent Processes.. R package version 1.6-4. 2014. URL http://cran.r-project.org/web/packages/lcmm/index.html

Rizopoulos D. JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data. Journal of Statistical Software. 2010; 35:1–33. [PubMed: 21603108]

Rizopoulos, D. JM: Joint Modeling of Longitudinal and Survival Data.. R package version 1.1-0. 2012. URL http://rwiki.sciviews.org/doku.php?id=packages:cran:jm

Rizopoulos, D. JMbayes: Joint Modeling of Longitudinal and Time-to-Event Data under a Bayesian Approach.. R package version 0.5-3. 2014. URL http://cran.r-project.org/web/packages/JMbayes/index.html

Schwarz G. Estimating the Dimension of a Model. The Annals of Statistics. 1978; 6:461–464.

Tsiatis AA, DeGruttola V, Wulfsohn MS. 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. 1995; 90:27–37.

Zhang D, Chen MH, Ibrahim JG, Boye ME, Wang P, Shen W. Assessing Model Fit in Joint Models of Longitudinal and Survival Data with Applications to Cancer Clinical Trials. Statistics in Medicine. 2014 doi:10.1002/sim.6269.

Figure 1.
An illustration of three bi-sectional partition methods.

Number of Subjects	
Subjects_in_lcss_long_anorexia	425
Subjects_in_lcss_surv_anorexia	425
Subjects_Used	425

Figure 2.
The number of subjects for pain.

Fit Statistics			
Log Likelihood	-6496.32		
AIC	13040.63	BIC	13137.88
AIC_{Long}	10867.68	BIC_{Long}	10920.35
$AIC_{Surv|Long}$	2172.96	$BIC_{Surv|Long}$	2217.53
$AIC_{Surv,0}$	2208.27	$BIC_{Surv,0}$	2248.79
ΔAIC	35.32	ΔBIC	31.27

Figure 3.
The fit statistics for pain under SPM1L with $J = 3$ based on LBSQP.

Survival Parameter Estimates (Survival Alone)							
Param/Var	Estimate	SE	DF	T-Value	P-Value	95% CI	Gradient
therapy	-0.4341	0.1025	425	-4.23	<.0001	(-0.636, -0.233)	-0.00322
race	0.1328	0.1898	425	0.70	0.4843	(-0.240, 0.506)	-0.00145
gender	0.08851	0.1356	425	0.65	0.5142	(-0.178, 0.355)	-0.00057
age	-0.06643	0.1059	425	-0.63	0.5309	(-0.275, 0.142)	-0.00183
karnofsky	-0.3209	0.1027	425	-3.13	0.0019	(-0.523, -0.119)	-0.0015
stage	-0.4745	0.1299	425	-3.65	0.0003	(-0.730, -0.219)	0.000437
bf	-0.08861	0.1140	425	-0.78	0.4375	(-0.313, 0.135)	-0.00052
$\log \lambda_1$	-1.6270	0.2444	425	-6.66	<.0001	(-2.107, -1.147)	0.001717
$\log \lambda_2$	-1.6475	0.2480	425	-6.64	<.0001	(-2.135, -1.160)	0.000158
$\log \lambda_3$	-1.1606	0.2398	425	-4.84	<.0001	(-1.632, -0.689)	-0.003

Figure 4.
The estimates of the parameters obtained by (11) with $J = 3$ based on LBSQP.

Parameter Estimates							
Param/Var	Estimate	SE	DF	T-Value	P-Value	95% CI	Gradient
Covariance Parameter Estimates							
Ω_{00}	0.6167	0.04796	423	12.86	<.0001	(0.522, 0.711)	0.192731
Ω_{10}	-0.03631	0.01373	423	-2.65	0.0085	(-0.063, -0.009)	-0.61855
Ω_{11}	0.06321	0.007785	423	8.12	<.0001	(0.048, 0.079)	-0.3081
σ	0.5404	0.005632	423	95.95	<.0001	(0.529, 0.551)	-0.03306
Longitudinal Parameter Estimates							
Intercept	0.6198	0.1795	423	3.45	0.0006	(0.267, 0.973)	0.036265
t	0.04090	0.01565	423	2.61	0.0093	(0.010, 0.072)	0.001122
therapy	-0.1145	0.07844	423	-1.46	0.1451	(-0.269, 0.040)	0.095573
race	-0.09783	0.1447	423	-0.68	0.4992	(-0.382, 0.187)	-0.03141
gender	-0.1813	0.1008	423	-1.80	0.0729	(-0.379, 0.017)	0.053837
age	-0.06076	0.08135	423	-0.75	0.4555	(-0.221, 0.099)	0.000031
karnofsky	-0.5782	0.07913	423	-7.31	<.0001	(-0.734, -0.423)	0.004468
stage	-0.08919	0.09446	423	-0.94	0.3456	(-0.275, 0.096)	0.013424
bf	0.1026	0.08937	423	1.15	0.2514	(-0.073, 0.278)	0.055028
Survival Parameter Estimates							
therapy	-0.4666	0.1075	423	-4.34	<.0001	(-0.678, -0.255)	0.039857
race	0.1553	0.1973	423	0.79	0.4317	(-0.233, 0.543)	-0.10528
gender	0.05964	0.1413	423	0.42	0.6732	(-0.218, 0.337)	-0.0278
age	-0.02925	0.1110	423	-0.26	0.7923	(-0.247, 0.189)	-0.1371
karnofsky	-0.4109	0.1089	423	-3.77	0.0002	(-0.625, -0.197)	0.068883
stage	-0.4191	0.1354	423	-3.10	0.0021	(-0.685, -0.153)	-0.01985
bf	-0.1133	0.1196	423	-0.95	0.3440	(-0.349, 0.122)	-0.0929
$\log \lambda_1$	-1.8586	0.2514	423	-7.39	<.0001	(-2.353, -1.364)	0.082103
$\log \lambda_2$	-1.8476	0.2549	423	-7.25	<.0001	(-2.349, -1.347)	0.027033
$\log \lambda_3$	-1.2576	0.2452	423	-5.13	<.0001	(-1.740, -0.776)	-0.06811
β	0.3816	0.06106	423	6.25	<.0001	(0.262, 0.502)	0.139638

Figure 5.
The estimates of the parameters for pain under SPM1L with $J = 3$ based on LBSQP.

Hazard Ratios & λ Estimates		
Parameter	**Estimate**	**95% CI**
HR_therapy	0.6272	(0.508, 0.775)
HR_race	1.1680	(0.793, 1.721)
HR_gender	1.0615	(0.804, 1.401)
HR_age	0.9712	(0.781, 1.208)
HR_karnofsky	0.6631	(0.535, 0.821)
HR_stage	0.6576	(0.504, 0.858)
HR_bf	0.8928	(0.706, 1.130)
λ_1	0.1559	(0.095, 0.256)
λ_2	0.1576	(0.096, 0.260)
λ_3	0.2843	(0.176, 0.460)
HR_β	1.4647	(1.299, 1.651)

Figure 6.
The hazard ratios and λ estimates for pain under SPM1L with $J = 3$ based on LBSQP.

(a) Anorexia

Fit Statistics			
Log Likelihood	-7079.46		
AIC	14206.92	BIC	14304.17
AIC_{Long}	12017.75	BIC_{Long}	12070.43
$AIC_{Surv|Long}$	2189.18	$BIC_{Surv|Long}$	2233.75
$AIC_{Surv,0}$	2208.27	$BIC_{Surv,0}$	2248.79
ΔAIC	19.10	ΔBIC	15.05

(a) Anorexia

(b) Cough

Fit Statistics			
Log Likelihood	-7201.63		
AIC	14451.26	BIC	14548.51
AIC_{Long}	12248.29	BIC_{Long}	12300.97
$AIC_{Surv|Long}$	2202.97	$BIC_{Surv|Long}$	2247.54
$AIC_{Surv,0}$	2208.27	$BIC_{Surv,0}$	2248.79
ΔAIC	5.31	ΔBIC	1.25

(b) Cough

(c) Dyspnea

Fit Statistics			
Log Likelihood	-6030.13		
AIC	12108.25	BIC	12205.50
AIC_{Long}	9911.77	BIC_{Long}	9964.45
$AIC_{Surv|Long}$	2196.48	$BIC_{Surv|Long}$	2241.05
$AIC_{Surv,0}$	2208.27	$BIC_{Surv,0}$	2248.79
ΔAIC	11.79	ΔBIC	7.74

(c) Dyspnea

(d) Fatigue

Fit Statistics			
Log Likelihood	-6569.90		
AIC	13187.80	BIC	13285.05
AIC_{Long}	11005.07	BIC_{Long}	11057.75
$AIC_{Surv|Long}$	2182.73	$BIC_{Surv|Long}$	2227.30
$AIC_{Surv,0}$	2208.27	$BIC_{Surv,0}$	2248.79
ΔAIC	25.54	ΔBIC	21.49

(d) Fatigue

Figure 7.
The fit statistics for anorexia, cough, dyspnea, and fatigue under SPM1L with $J = 3$ based on LBSQP.

Fit Statistics			
Log Likelihood	-4051.49		
AIC	8148.99	BIC	8240.79
AIC_{Long}	6104.63	BIC_{Long}	6156.52
$AIC_{Surv|Long}$	2044.36	$BIC_{Surv|Long}$	2084.27
$AIC_{Surv,0}$	2075.71	$BIC_{Surv,0}$	2107.64
ΔAIC	31.35	ΔBIC	23.37

(b) $D_{Long} + D_{Surv}$

Fit Statistics			
Log Likelihood	-4079.49		
AIC	8204.98	BIC	8296.78
AIC_{Long}	6151.87	BIC_{Long}	6203.76
$AIC_{Surv|Long}$	2053.11	$BIC_{Surv|Long}$	2093.03
$AIC_{Surv,0}$	2075.71	$BIC_{Surv,0}$	2107.64
ΔAIC	22.60	ΔBIC	14.61

(c) $D_{Long1} + D_{Surv}$

Fit Statistics			
Log Likelihood	-4399.92		
AIC	8845.84	BIC	8937.64
AIC_{Long}	6778.37	BIC_{Long}	6830.26
$AIC_{Surv|Long}$	2067.47	$BIC_{Surv|Long}$	2107.39
$AIC_{Surv,0}$	2075.71	$BIC_{Surv,0}$	2107.64
ΔAIC	8.24	ΔBIC	0.25

(d) $D_{Long2} + D_{Surv}$

Fit Statistics			
Log Likelihood	-4775.45		
AIC	9596.91	BIC	9688.71
AIC_{Long}	7525.67	BIC_{Long}	7577.56
$AIC_{Surv|Long}$	2071.23	$BIC_{Surv|Long}$	2111.15
$AIC_{Surv,0}$	2075.71	$BIC_{Surv,0}$	2107.64
ΔAIC	4.47	ΔBIC	-3.51

(e) $D_{Long3} + D_{Surv}$

Figure 8.
The fit statistics for $D_{Long} + D_{Surv}$, $D_{Long1} + D_{Surv}$, $D_{Long2} + D_{Surv}$, and $D_{Long3} + D_{Surv}$ under SPM2L.

Table 1

AIC$_{Surv|Long}$'s and AIC's for pain under SPM1L with different partition algorithms and different J's.

	AIC$_{Surv\|Long}$			AIC		
J	**LBSQP**	**MBSQP**	**RBSQP**	**LBSQP**	**MBSQP**	**RBSQP**
1	2200.94	2200.94	2200.94	24.87	24.87	24.87
2	2170.94	2170.94	2170.94	35.35	35.35	35.35
3	2172.96	2172.96	2172.90	35.32	35.32	34.84
4	2174.91	2174.91	2174.91	34.81	34.81	34.81
5	2176.53	2174.36	2176.84	34.66	35.04	34.68
6	2175.98	2173.66	2176.15	34.89	35.72	35.36
7	2175.29	2175.29	2175.60	35.57	35.57	35.59
8	2177.22	2177.22	2177.22	35.44	35.44	35.44

Table 2

AIC$_{\text{Surv|Long}}$'s for cough and pain under SPM1L with $t_{\max,i}$ adjustments using different w's.

	Cough		Pain	
w	TMAXI=1	TMAXI=2	TMAXI=1	TMAXI=2
0	2200.97	2200.88	2170.94	2168.33
0.25	2200.48	2200.74	2172.84	2171.65
0.5	2201.71	2201.51	2177.09	2175.40
0.75	2204.21	2203.63	2182.34	2180.72
1[a]	2206.68	2206.68	2186.46	2186.46

[a] No $t_{\max,i}$ adjustment