%JM Macro
Reference Manual

Table of Contents

1. Introduction

Longitudinal studies are often used in clinical research to investigate disease progression and its relationship with a number of risk factors or health outcomes. In these studies, different types of endpoints are collected, such as the time when certain events appear or repeated measurements of biomarkers. In occasions, interest may lie in questions that require the joint modelling of some of these endpoints. For further background on joint modeling of longitudinal and time-to-event data, see the reviews of Tsiatis and Davidian [1] and Wu et al [2] and Rizopoulos' book [3].

Three R packages, *JM* [5, 6], *joineR* [7] and *lcmm* [8] and the *stjm* Stata command [9] fit these models using maximum likelihood whilst the *JMBayes* R package [10] uses Markov chain Monte Carlo simulations to solve this problem from a Bayesian perspective.

The %JM macro has been developed to fit a broad range of models in this family. Four possible distributions are available for the longitudinal response (normal, binary, Binomial, Poisson). A range of options are available to fit subjects' evolutions of the longitudinal response over time (a linear function, splines, natural cubic splines and B-splines). For the time-to-event response, the %JM macro fits parametric, stratified or not, proportional hazard models. The baseline hazard function needs to be specified with one of the following options: exponential, Weibull, piecewise exponential, generalizations of Gompetz and Weibull model based on splines and the Royston-Parmar model. The macro offers multiple options to connect the time-to-event model and the longitudinal model: a (trajectory) current-value-dependent shared parameter, a (trajectory) slope-dependent shared parameter, lagged effects, cumulative effects, random effects coefficients and interaction effects.

2. Joint modeling of longitudinal and time-to-event data

The %JM macro fits a group of latent variable models where the association structure between outcomes is modeled via latent variables. The macro uses normally-distributed random effects b_i to account the association between the longitudinal and the event models. Given the random effects b_i, the observations in the longitudinal response y_i and the time-to-event responses T_i are assumed to be independent and thus expressed as

$$p(T_i, \delta_i, y_i \mid b_i; \theta) = p(T_i, \delta_i \mid b_i; \theta)\, p(y_i \mid b_i; \theta) \qquad [2.1]$$

where T_i is the observed time-to-event or time-to-censoring, δ_i is the indicator of event or censoring, $y_i = \{y_{ij} ; 1 \leq j \leq n_i\}$ the collection of observations in the longitudinal response for subject i and θ the vector of parameters that includes parameters for the time-to-event response θ_T and for the longitudinal model θ_y. Under the assumptions of this model, the marginal (joint) log-likelihood of the observations on subject i is thus expressed as follows:

$$p(T_i, \delta_i, y_i \mid \theta_y, \theta_T) = \int p(T_i, \delta_i \mid b_i; \theta_y, \theta_T) \left[\prod_j p(y_i(t_{ij}) \mid b_i; \theta_y) \right] p(b_i; \theta_b)\, db_i \qquad [2.2]$$

where θ_b denotes the unique parameters of the random effects covariance matrix G.

The %JM macro fits random-effects (G-sided, conditional) mixed-effects models as opposed to the covariance-patter (R-sided, marginal) mixed-effects models where the correlations among the residuals are fitted using different structures for the R matrix. In this context, example 38.5 of SAS documentation [11] illustrates the use of Proc GLIMMIX to fit a joint model of two distinct distributions using either the G or the R matrix. Nevertheless, GLIMMIX does neither handle censored time-to-event data nor fit proportional hazards models; features covered by the %JM macro.

2.1. The longitudinal submodel

Following the general conditional (G-sided) mixed models framework, the macro uses latent random effects vectors b_i, independent across subjects, to model the within-subject covariance. Given the random effects b_i the observations for the longitudinal response on the same subject are assumed to be independent.

$$p(y_i \mid b_i; \theta_y) = \prod_j p(y_i(t_{ij}) \mid b_i; \theta_y) \qquad [2.3]$$

More specifically, we will express this generalized linear mixed effects model as:

$$g\{m_i(t)\} = g[E\{y_i(t) \mid b_i\}] = X_i(t)(\beta_t + b_i) + Z_i \beta_b \qquad [2.4]$$

where $m_i(t)$ denotes the expected value for subject i at time t on the longitudinal response $y_i(t)$, $g(\cdot)$ is a link function, $X_i(t)$ is the design matrix of fixed and random effects that model the trajectories of the longitudinal response over time, associated with the coefficients β_t and b_i respectively, and Z_i are β_b are the design matrix and the coefficients associated with the baseline covariates.

The %JM macro fits either normal, binary, binomial or Poisson longitudinal responses by specifying the distribution in the LongiType macro parameter. The user might change the link function, from the canonical link that is used by default, by specifying LongiLink = identity | logit | log | probit. The longitudinal model can include one or more baseline covariates (LongiCovariates=<variables list>).

%JM does not exclude any longitudinal measures. If one wishes to exclude the longitudinal measures after the survival times, those values should be excluded in the input dataset. Next section explains how the macro fits trajectories of the longitudinal response over time.

2.1.1. Modeling trajectories over time

Random slopes

The %JM macro can fit a simple random intercepts and slopes model with a fixed-effect intercept B_0 and a fixed-effect slope B_1 as well as a random-effect intercept b_{i0} and a random-effect slope b_{i1} (LongiTimeModel= LINEAR). The longitudinal model can also include variables that interact with the time slope (LongiTimeInteraction= <variables list>) as illustrated the model given below where LongiTimeInteraction= TRT was used.

$$g\{m_i(\text{t})\} = \beta_0 + \beta_1 t + \beta_2 GENDER_i + \beta_3 t\, TRT_i + b_{i0} + b_{i1}t, \quad (b_{i0}, b_{i1}) \sim \text{N}\begin{bmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{bmatrix} \quad [2.5]$$

Section 2.1.2 details the options that %JM provides to model the correlation between the random effects b_{i0} and b_{i1}.

Splines

A spline is a function that is constructed piecewise from polynomials leading to a fully parametric but flexible approach to fit non-linear subject specific trajectories in the longitudinal response. By using LongiTimeModel= SPLINE the SAS macro defines a linear combination of truncated power functions of time as shown in equation 2.6 for a piecewise cubic (LongiDegree=3) polynomial with N internal knots (LongiNknots=<number>).

$$g\{m_i(t)\} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t + (\beta_2 + b_{i2})t^2 + (\beta_3 + b_{i3})t^3 + \quad [2.6]$$

$$+(\beta_{3+1} + b_{i,3+1})(t - \upsilon_1)_+^3 + \cdots + (\beta_{3+N} + b_{i,3+N})(t - \upsilon_N)_+^3 + Z_i\beta_b$$

where the expression $(\cdot)_+$ is used to indicate that negative values are set to zero, υ_1 to υ_N are N internal knots, and Z_i are β_b are the design matrix and the coefficients associated with the baseline covariates. If

internal knots are not given by the user (Longiknot1=<time>, Longiknot2=<time>, etc), %JM calculates the knots in equally spaced quantile partition (ESQP) of the longitudinal-response observation times.

The above expression can be generalized in terms of basis functions $B_j(x)$ as given below

$$g\{m_i(t)\} = \sum_{j=0}^{N+n}(\beta_j + b_{ij})B_j(t) + Z_i\beta_b \tag{2.7}$$

where N is the number of internal knots, n the degree, and $B_0(x)=1$, $B_0(x)=t$, $B_2(x)=t^2$, … , $B_n(x)=t^n$, $B_{n+1}(x)=(t-\upsilon_1)^1_+$, … , $B_{n+N}(x)=(t-\upsilon_N)^n_+$ are N+n+1 basis function constructed from truncated power functions.

B-splines

Any piecewise polynomial function can be expressed as a linear combination of Bézier basis functions $B_{i,n}(x)$ through the generalization of the Bézier curve. B-splines are maximally differentiable bases with a number of desirable properties that in general lead to better convergence as compared to the bases constructed from truncated power functions [12, 13]. Using LongiTimeModel= BSPLINE, %JM produces the basis functions $B_j(x)$ using the Boor recurrence relation [14].

For example, for a quadratic (LongiDegree=2) B-spline with two internal knots (LongiNKnots=2) placed at the 33.3% and 66.7% percentiles of the follow-up times (partition used by default if the user does not provide the knots) and a variable that interact with the time slope (LongiTimeInteraction= TRT) the %JM macro creates five basis functions variables $B_0(x)$ to $B_4(x)$[1]. This model requires nine fixed-effect parameters, β_0 to β_8, and five random-effects parameters, b_{i0} to b_{i4}, to fit the evolution of the longitudinal responses over time as shown below,

$$g\{m_i(t)\} = \sum_{j=0}^{4}(\beta_j + b_{ij})B_j(t) + \sum_{j=1}^{4}\beta_{4+j}B_j(t)TRT_i + Z_i\beta_b \tag{2.8}$$

B-splines basis functions are set to zero beyond the boundary knots so this model fits a horizontal line beyond them. If not given by the user, the boundary knots are fixed as 0 and the maximum time of the longitudinal responses respectively.

---

[1] The first basis function $B_0(x)$ obtained using the Boor recurrence relation is created and added to the dataset but not used. The %JM macro uses instead the intercept term in the model, what is equivalent to setting the first basis $B_0(t)$ to 1 at any point t. This approach is similar to the use of the default option Intercept=False in the bs() function of R.

Natural cubic splines

Natural cubic splines (LongiTimeModel = NATURALCUBIC) are cubic splines resulting after setting the second and third derivatives to zero at the boundary knots and therefore fitting a straight line outside these. Four (two fixed and two random) less parameters are required with respect to the complete cubic spline with the same number of internal knots. The basis functions $B_0(t)$ to $B_{N+1}(t)$ calculated by the %JM macro using the following equation:

$$B_0(t) = 1, \quad B_1(t) = t, \quad B_{j+1}(t) = d_j(t) - d_N(t) \qquad [2.9]$$

where N is the number of internal knots, $d_j(t) = (t - \upsilon_j)_+^3 / (\upsilon_{N+1} - \upsilon_j)^3$ for j indexed from 1 to N, υ_0 is the lower boundary knot, υ_1 to υ_N are N internal knots and υ_{N+1} refers to the upper boundary knot. If not given by the user, υ_0 is set to zero and υ_{N+1} to the maximum time of the longitudinal responses.

2.1.2.   Random effects variance-covariance structures

The macro allows three possible structures for the G matrix. Using an unstructured G matrix (LongiGMatrix=UN), k(k+1)/2 covariance parameters are required where k is the number of parameters required to fit the longitudinal trajectories. For example if k=3, six parameters are required to define G, as illustrated below.

$$G = \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{0k} & \sigma_{12} & \sigma_2^2 \end{bmatrix} \qquad [2.10]$$

With a variance-components option (LongiGMatrix=VC), the G matrix is diagonal and only k variance parameters are required.

$$G = \begin{bmatrix} \sigma_0^2 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix} \qquad [2.11]$$

If an ante-dependence matrix is chosen (LongiGMatrix=ANTE(1)), 2k-1 variance parameters are needed.

$$G = \begin{bmatrix} \sigma_0^2 & \rho_{01}\sigma_0\sigma_1 & \rho_{01}\rho_{12}\sigma_1\sigma_2 \\ \rho_{01}\sigma_0\sigma_1 & \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{01}\rho_{12}\sigma_1\sigma_2 & \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \qquad [2.12]$$

2.1.3.   Parameterizations of the G matrix

This macro allows three possible parameterizations for G matrix as illustrated in Table 1 for an example with 3 parameters. By default, LongiGMatrixParam = VARCOVAR is used and parameters are created for (log-transformed square root of) variances and covariances. Using LongiGMatrixParam = CHOL, the macro creates parameters for the Choleski decomposition (C) of G [G = CC'].

When an ante-dependence G matrix is used LongiGMatrixParam needs to be set to VARCORR so the macro defines parameters for (log-transformed square root of) variances and correlation coefficients between adjacent parameters.

Table 1. Parameterizations of the G matrix in NLMIXED

| Variance Component | LongiGMatrix=UN or VC | | LongiGMatrix=ANTE(1) |
| | LongiGMatrixParam= | | LongiGMatrixParam= |
	VARCOVAR	CHOL	VARCORR
σ_0^2	EXP(2*L_LOGSD_INTERCEPT)	L_GCHOL_B_1_1*L_GCHOL_B_1_1	EXP(2*L_LOGSD_INTERCEPT)
σ_{01}	L_COV_B0_B1	L_GCHOL_B_2_1*L_GCHOL_B_1_1	EXP(L_LOGSD_INTERCEPT+L_LOGSD_TIME_B1)*L_RHO_B0_B1
σ_1^2	EXP(2*L_LOGSD_TIME_B1)	L_GCHOL_B_2_1*L_GCHOL_B_2_1+ L_GCHOL_B_2_2*L_GCHOL_B_2_2	EXP(2*L_LOGSD_TIME_B1)
σ_{02}	L_COV_B0_B2	L_GCHOL_B_3_1*L_GCHOL_B_1_1	EXP(L_LOGSD_INTERCEPT+L_LOGSD_TIME_B2)*L_RHO_B0_B2
σ_{12}	L_COV_B1_B2	L_GCHOL_B_3_1*L_GCHOL_B_2_1+ L_GCHOL_B_3_2*L_GCHOL_B_2_2	EXP(L_LOGSD_TIME_B1+L_LOGSD_TIME_B2)*L_RHO_B1_B2
σ_2^2	EXP(2*L_LOGSD_TIME_B2)	L_GCHOL_B_3_1*L_GCHOL_B_3_1+ L_GCHOL_B_3_2*L_GCHOL_B_3_2+ L_GCHOL_B_3_3*L_GCHOL_B_3_3	EXP(2*L_LOGSD_TIME_B2)

2.2.   The time-to-event submodel

The %JM macro fits parametric proportional hazards models in which the baseline hazard function needs to be specified. The variable of interest is the time to the event T_i^* that might not be observed for all subjects. Two variables are observed: time to event or censoring T_i (TimeVar = <Variable Name>) and the indicator of event $\delta_i = I(T_i^* \leq C_i)$ (EventVar=<Variable Name> and EventVal = <Value>).

The general expression of the time-to-event model is given by

$$h_i(t) = h_0(t) \exp(\gamma^T \omega_i) \qquad [2.13]$$

where $\omega_i^T = (\omega_{i1}, \ldots \omega_{i1})$ denotes the vector of covariates (EventCovariates=<Variables list>) and γ denotes de corresponding vector of coefficients.

2.2.1.   Specifying the baseline hazard

The %JM macro offers several options for specifying the baseline hazard $h_0(\cdot)$:

- Exponential
- Weibull
- Piecewise exponential
- Generalization of the Gompertz model based on splines
- Generalization of the Weibull model based on splines
- Royston-Parmar model

Exponential

The simplest model for the time-to-event response is to assume that the baseline risk is constant over time (EventModel= EXPONENTIAL):

$$h_0(t) = \lambda \qquad\qquad [2.14]$$

Weibull

The Weibull distribution adds a shape parameter to allow the hazard risk to increase or decrease over time using a power function of time (EventModel= WEIBULL).

$$h_0(t) = \lambda^k k t^{k-1} \qquad\qquad [2.15]$$

Piecewise exponential

In standard survival analysis it is frequent to leave $h_0(\cdot)$ unspecified as in Cox model. In the joint modeling framework such approach may underestimate the standard errors of the parameters estimates [50]. With EventModel= PIECEWISE the macro uses a piecewise-constant hazard model with a hazard function defined as follows:

$$h_0(t) = \sum_{q=1}^{Q+1} \xi_q I(\upsilon_{q-1} < t \le \upsilon_q) \qquad\qquad [2.16]$$

where $\upsilon_0 = 0, \upsilon_1, \dots, \upsilon_{Q-1}$ are Q internal knots that split the range of times into Q intervals, υ_{Q+1} is a value higher than the maximum observed time, and ξ_q is the hazard at the interval defined from υ_{q-1} to υ_q. The %JM macro allows the user to fix the number (EventNknots= <n>) and the position of these knots (Eventknot1= <t1>, Eventknot1= <t2>). If not specified by the user, six knots are used that split the time period into seven intervals. The position of the internal knots is either given by the user or placed automatically by the macro in an equally spaced quantile partition of event times.

In order to calculate the partition, the user might choose either to use the default option, all censored or non-censored event times or to change this and use only actual (non-censored) events to calculate that partition (AdditionalOptions = EVENTTIMES).

The piecewise exponential model can get as close to the unspecified Cox model as desired just by increasing the number of knots. However, too many knots are not advisable in the joint model framework since the standard errors of the parameter estimates might be underestimated [15].

Generalizations of the Gompertz model

The Gompertz model can be formulated in terms of the baseline log-hazard function as $\log[h_0(t)] = k_0 + k_1 t$. A flexible generalization of this model can be constructed as $\log[h_0(t)] = s(t; \kappa)$ where $s(t; \kappa)$ is a smooth function of t, with κ the vector of coefficients of the smooth function (EventModel= GSPLINE|GBSPLINE|GNATURALCUBIC). For example, a generalization using natural cubic splines (EventModel= GNATURALCUBIC) has to following formulation

$$\log\{h_0(t)\} = \sum_{j=0}^{Q+1} \kappa_j B_j(t) \qquad\qquad [2.17]$$

where Q is the number of internal knots and $B_0(t)$ to $B_{Q+1}(t)$ are the basis functions defined in section 2.1.1.

Generalizations of the Weibull model

The Weibull model is formulated in terms of the baseline log-hazard function as $\log[h_0(t)] = k_0 + k_1\log(t)$. A flexible generalization of this model can be constructed as $\log[h_0(t)] = s(x; \kappa)$ where $s(x; \kappa)$ is a smooth function of x, with $x=\log(t)$, and κ the vector of coefficients of the smooth function (EventModel= WSPLINE|WBSPLINE|WNATURALCUBIC).

Royston-Parmar model

The Weibull model can also formulated in terms of the log-transformed baseline cumulative log-hazard function[2] as a monotonic linear equation $\log\{H_0(t)\} = k_0 + k_1\log(t)$ with $\kappa_1>0$. Royston and Parmar proposed a flexible generalization of this model as $\log[H_0(t)] = s(x; \kappa)$ where $s(x; \kappa)$ is a smooth function of x, with $x=\log(t)$ and κ the vector of coefficients of the smooth function [16] (EventModel= RPSPLINE|RPBSPLINE|RPNATURALCUBIC).

---

[2] $H_0(t) = \int_0^t h_0(s)ds$

2.2.2.   Stratification

In occasions, it is not reasonable to assume that the whole sample has the same baseline hazard function, that is, the proportional hazard assumption between subgroups do not hold. In these situations it is convenient to treat those subgroups as strata where a different baseline hazard is fitted for each of the strata (EventStrata= <variable>), as given by

$$h_{ik}(t) = h_{0k}(t) \exp\{\gamma^T \omega_i\}$$ [2.18]

where $h_{0k}(t)$ would be the baseline risk of stratum k. This option multiplies the number of required parameters, as explained in previous sections, by the number of strata.

2.2.3.   Exogenous time-dependent covariates

One of the motivations for the development of joint models for longitudinal and time-to-event responses was to find an alternative to the extended Cox model, also known as the Andersen-Gill model, when the time-dependant covariate $y_i(t)$ is not exogenous. Endogenous time-dependant covariates are thus included in this model as the dependant longitudinal response $y_i(t)$. However there could be situations where, in addition to the endogenous time-dependant response $y_i(t)$, it is of interest to include a exogenous time-dependant covariate $w_i(t)$ following the principles of the extended Cox model. In this case, the general expression of the time-to-event model changes to

$$h_i(t) = h_0(t) \exp\{\gamma_i^T \omega_i(t)\}$$ [2.19]

where $\omega_i(t)$ is a covariate vector that contains both baseline and exogenous time-dependant covariates. Exogenous time-dependent-covariates are included in the macro using a vertical structure (see section 7.2).

2.3.   Shared parameters

2.3.1.   A current-value dependent shared parameter

The most standard approach to connect the longitudinal model and the time-to-event model in the random-effects shared-parameter models framework is to assume that the risk of event at a given time t depends on the estimated value of the longitudinal response at that time (SharedParam= CURRENT_VALUE) resulting in the following equation:

$$h_i(t) = h_0(t) \exp\{\gamma^T \omega_i + \alpha \, m_i(t)\}$$ [2.20]

For the case that the baseline risk $h_0(t)$ is set as exponential, it should be noted that, once a shared-parameter α is included between the subject-level estimate of the longitudinal response $m_i(t)$ and the time-to-event response T_i, the resulting hazard $h_i(t)$ for a subject i given by equation 2.20 is not constant anymore but changes over time if the longitudinal response changes over time. Thus in this setting, the exponential baseline might be a reasonable choice in situations where the change of risk over time is expected to be mostly explained through the changes in the longitudinal response $m_i(t)$.

We can decompose the longitudinal response at t as the sum of the baseline value and the change from baseline as given by $m_i(t) = m_i(0) + ch\_m_i(t)$ where $ch\_m_i(t) = m_i(t) - m_i(0)$.

From %JM macro version 02.01 it is possible to fit separately the effect of the baseline value and the effect of the change during the study on the log-risk function as given by

$$h_i(t) = h_0(t)\exp\{\gamma^T\,\omega_i + \alpha_1\,m_i(0) + \alpha_2\,ch\_m_i(t)\} \qquad \text{[2.20bis]}$$

where $ch\_m_i(t) = m_i(t) - m_i(0)$ using SharedParam= BASELINE_VALUE CHANGE_FROM_BASELINE.

2.3.2. A slope dependent shared parameter

For each patient, the longitudinal model fits a trajectory in time and it may be the case that it is not the value but the rate of rise (or decrease) of the longitudinal response what increases the risk of event (SharedParam= SLOPE).

$$h_i(t) = h_0(t)\exp\{\gamma^T\,\omega_i + \alpha\,m_i'(t)\} \qquad \text{[2.21]}$$

2.3.3. Cumulative effects

This options should be used when the event is not expected to depend on specific time points of the longitudinal response but on the cumulative effect (SharedParam= CUMULATIVE).

$$h_i(t) = h_0(t)\exp[\gamma^T\,\omega_i + \alpha\,g\{m_i(t)\}] \qquad \text{[2.22]}$$

2.3.4. Lagged effects

In occasions there is a latency period between the value of the longitudinal response and the appearance of an increased risk. These situations can be modeled fixing a time-lagged effect model. The model below is fitted for a time-depend parameterization if lagging is specified (SharedLagging = <time lag>).

$$h_i(t) = h_0(t)\,\exp\{\gamma^T\,\omega_i + \alpha\,m_i[\max(t-c,0)]\} \qquad [2.23]$$

where c specifies the time lag of interest.

Lagging can be combined with other parameterizations. Using, for instance, SharedParam= CUMULATIVE and SharedLagging=<period> would result in equation 2.24.

$$h_i(t) = h_0(t)\,\exp\left\{\gamma^T\,\omega_i + \alpha \int_0^{\max(t-c,0)} m_i(s)\,ds\right\} \qquad [2.24]$$

2.3.5.   Random-effects coefficients

%JM also allows connecting the longitudinal model and the time-to-event model through isolated random effects of the longitudinal (SharedParam= COEFFICIENTS). If for example a random intercepts and slopes model is used for the longitudinal response the following model can be fitted:

$$h_i(t) = h_0(t)\,\exp\{\gamma^T\,\omega_i + \alpha_1\,b_{i0} + \alpha_2\,b_{i2}\} \qquad [2.25]$$

The user needs to specify what random effects need to be included. The example of equation 2.25 would be fitted by specifying SharedCoefficients = bi0 bi1.

2.3.6.   Multiple shared-parameters

Several parameterizations, separated between blank spaces in the SharedParam parameter, can combined leading to multiple shared parameters α_i. For example, SharedParam= CURRENT_VALUE SLOPE would lead to the following model

$$h_i(t) = h_0(t)\,\exp\{\gamma^T\,\omega_i + \alpha_1\,m_i(t) + \alpha_2\,m_i'(t)\} \qquad [2.26]$$

Including multiple parameterizations should be done cautiously. For example it is possible to include both a time-dependant shared-parameter α_1 and a cumulative-dependant parameters α_2, however such a model might have high collinearility between both effects.

2.3.7.   Interaction effects

In occasions it is not correct to assume that the effect of the longitudinal response $m_i(t)$ on the event risk $h_i(t)$ is the same for all subjects and shared-parameter interactions need to be fitted (SharedInteraction=<var list>).

For example, SharedInteraction=TRT would change the basic time-dependant model of section 2.3.1 into the following model:

$$h_i(t) = h_0(t)\exp\{\gamma^T\,\omega_i + \alpha_1 m_i(t) + \alpha_2 m_i(t)\cdot TRT\} \qquad [2.27]$$

2.3.8.   Association term

For non-normal longitudinal responses, the link function is usually distinct from the identity function leading to two options to link the longitudinal and the time-to-event models: a) a linear relationship between the log-hazard function and the expected response of the longitudinal response $m_i(t)$ as shown in equation 2.20 (SharedLongiTerm= MEAN); and b) the log-hazard function linearly related to the linear predictor $g\{m_i(t)\} = X_i(t)$ + $Z_i(t)b_i$ as illustrated in the equation below (SharedLongiTerm=LINEAR_PREDICTOR).

$$h_i(t) = h_0(t)\exp[\gamma^T\,\omega_i + \alpha\,g\{m_i(t)\}] = h_0(t)\exp[\gamma^T\,\omega_i + \alpha\,X_i(t)\big(\beta_t + b_i\big) + Z_i\beta_b] \quad [2.28]$$

where $m_i(t)$ is the estimated value for longitudinal response $y_i(t)$ and $g(\cdot)$ the link function.

2.4.   Obtaining the maximum-likelihood estimates

The %JM macro uses the NLMIXED procedure to fit the joint model. NLMIXED provides two useful features in this context:

a) It allows to manually define the (log-transformed) probability function conditional to the random effects, that is different for the longitudinal response and the time-to-event response, through the use of the LL statement (see section 2.4.1).

b) By using the RANDOM statement, the NLMIXED procedure takes care of the calculation of the marginal likelihood by integrating the conditional probability function over the random effects (see section 2.4.2).

2.4.1. Conditional likelihood function

For the longitudinal response, the log-transformed probability function conditional to the random effects parameters has always a closed form. For example, for a normal longitudinal response this is defined as follows

$$\log\{p\big(y_i(t_{ij})|\, b_i;\theta_y\big)\} = -\frac{\log(2\pi)}{2} - \log(\sigma_e) - \frac{1}{2}\frac{\{y_i(t_{ij}) - m_i(t_{ij})\}^2}{\sigma_e^2}$$ [2.29]

where $\theta = (\beta_t,\ \beta_b,\ \sigma^2)$.

On the other hand, the log-transformed probability function of the time-to-event model, conditional to the random effects parameters, is expressed as follows

$$\log\{p(T_i,\delta_i|\, b_i;\theta,\beta)\} = \delta_i \log\{h_i(T_i|\, b_i;\theta,\beta)\} - \int_0^{T_i} h_i(t|\, b_i;\theta,\beta)dt$$ [2.30]

In order to calculate this, the macro needs to approximate the integral involved in the second part of the above equation.

Approximation of the survival function (standard data)

This integral from 0 to T_i of the hazard function conditioned to the random effects might or might not have a close form depending on the model chosen for both the longitudinal and the time-to-event responses. The %JM macro approximates this integral using the 15-points Gauss-Kronrod rule as follows:

$$\int_0^{T_i} h_i(t|\, b_i;\theta,\beta)dt \approx \frac{T_i}{2}\sum_{k=1}^{15} w_k h_i(t_{ik}|\, b_i;\theta,\beta)$$ [2.31]

where t_{ik} are the 15 Kronrod-rule nodes for the integral from −1 to 1 re-scaled into the 0 to T_i interval, and w_k are the 15 Kronrod-rule weights for the integral from -1 to 1 respectively.

If splines are used to model $m_i(t)$, multiple variables need to be added to the work dataset. For example, in a model with a 5-parameter spline the macro adds 5 {bases $B_j(T_i)$} + 15 {nodes t_{ik}} + 15 {weights w_k}+ 15*5 {bases $B_j(t_{ik})$} = 110 variables in the work dataset.

Approximation of the survival function (vertical format)

When a vertical data structured is used (DataStructure = VERTICAL) the input dataset needs to have the form as shown in Table 2 (see section 7.2).

Table 2. A vertical data structure

SubjectID i	TimeFrom $T_{i,j-1}$	TimeTo $T_{i,j}$	EventYN δ_{ij} [3]
1	0	2	0
1	2	3	0
1	3	10	0
1	10	12	1

The log-transformed probability function of a time T_i can be decomposed as the following sum:

$$\mathrm{LL} = \sum_{j=1}^{P}\left[\delta_{ij}\log\{h_i(T_i|\, b_i;\theta,\beta)\} - \int_{T_{i,j-1}}^{T_{i,j}} h_i(t|\, b_i;\theta,\beta)\mathrm{d}t\right] \qquad [2.32]$$

where $T_{i,1}$ to $T_{i,P-1}$ are the P-1 internal knots used to split the time-to-event response of subject *i* into a long (vertical) format, $T_{i0} = 0$ and $T_{i,P}=T_i$. In these situations the Kronrod rule is applied to each interval leading to both a more accurate approximation of this integral and an increase in the time required for this operation.

2.4.2. Marginal likelihood function

Under the assumptions of this model, the marginal (joint) log-likelihood of the observations on subject *i* can be expressed as follows:

$$\log \mathrm{p}(T_i,\delta_i,y_i\,|\,\theta_y,\theta_T) = \log \int \mathrm{p}(T_i,\delta_i|\, b_i;\theta_y,\theta_T)\left[\prod_j \mathrm{p}(y_i(t_{ij})|\, b_i;\theta_y)\right]\mathrm{p}(b_i;\theta_b)\mathrm{d}b_i \qquad [2.33]$$

where T_i is the observed time-to-event or time-to-censoring, δ_i is the indicator of event or censoring, $y_i(t_{ij})$ the collection of observations in the longitudinal response for subject *i*, θ_T and θ_y the vectors of parameters for the time-to-event and the longitudinal model respectively, and θ_b denotes the unique parameters of the random effects covariance matrix G.

The NLMIXED procedure approximates the marginal likelihood, the integral of the conditional probability function over the random effects, using adaptive or non-adaptive Gauss-Hermite quadrature. By default, the NLMIXED procedure uses adaptive Gauss-Hermite quadrature, and selects the number of quadrature points adaptively by evaluating the log-likelihood function at the initial parameters until two successive evaluations have a relative difference less than 0.0001. The NLMIXEDOptions macro

---

[3] $\delta_{ij} = \delta_i I(j = P)$ and P is the last time interval.

parameter allows personalization of the approximation method with the options available in the NLMIXED procedure. For example, NLMIXEDOptions=noad noadscale qpoints=100 would lead to a non-adaptive Gauss-Hermite quadrature with 100 quadrature points.

2.4.3. Optimization routines

%JM allows using of the optimization routines available in the NLMIXED procedure through the NLMIXEDOptions macro parameter. If this macro parameter is not provided, the default NLMIXED options are used. Table 3 shows the execution times that were required to fit the joint model over the aids dataset illustrated in section 6.1.

Table 3. Execution Times Using different Optimization Techniques[4]

Optimization Technique	NLMIXEDOptions	Time (minutes)	-2LL
Dual Quasi-Newton	tech=quanew (or left blank)	11:36	8656.8
Double-dogleg	tech=dbldog	8:04	8656.8
Newton-Raphson with search	tech=newrap	17:28	8656.8
Newton-Raphson with ridging	tech=nrridg	17:42	8656.8
Trust region	tech=trureg	17:32	8656.8

The convergence criterion of NLMIXED is based on the relative change in the gradient changes between two iterations. In occasions the relative changes are very small, and thus triggering the stopping of the algorithm, but the values themselves are not that small. It has been observed that in the context of joint models, in occasions the convergence criterion is met but the maximum absolute gradient is still > 0.01 and even >0.1 leading to poor parameter estimates. In these situations the SAS system gives the following: "NOTE: At least one element of the (projected) gradient is greater than 1e-3". This issue can be solved by setting NLMIXEDOptions = GCONV=0 to ensure that convergence is not achieved based on the relative gradient and therefore either the absolute gradient or the relative function is required to meet the convergence criterion.

2.4.4. Initial parameters

The %JM macro fits first the longitudinal model and the parametric time-to-event model separately, using either the MIXED procedure or the GLIMMIX procedure for normal and non-normal longitudinal responses respectively and the LIFEEG procedure for the time-to-event responses.

---

[4] In all these models NLMIXEDOptions = QPOINTS = 5 GCONV=0 was used.

The estimates obtained from these disjoint models, together with zeros for the association parameters α, are provided as initial parameters to the NLMIXED procedure to fit the joint model. Alternatively, users can provide initial parameters, using InitialParameters=<dataset>, that have previously stored in a dataset. Section 7.3 specifies the structure expected for the dataset of initial parameters. Sometimes it is useful to fit an initial model, save the parameters estimates in a dataset using the OutputParameters macro parameter, and then use these estimates as initial parameters in subsequent calls to the macro.

For time-to-events models using splines, initial parameters are set using a simplified model where the baseline hazard is assumed constant over time.

3.   Compiling and using the macro

3.1.   Compiling the macro

The following six macros need to be compiled for the %JM macro to work:

```
%include '&location./calculateknotspartition.sas';
%include '&location./spline.sas';
%include '&location./ncspline.sas';
%include '&location./bspline.sas';
%include '&location./kronrodrule15p.sas';
%include '&location./jm.sas';
```

Session compiled macros exist only during the current SAS session. A convenient solution to avoid the %INCLUDE statements above in each session is to include these macros in a location referred by the SAS AUTOCALL facility.

3.2.   Calling the macro

The %JM macro needs to be called using keyword parameters. For example, the joint model illustrated in section 6.1 is fitted using the following call to the %JM macro.

```
%JM(        Data                = AIDS,
            SubjectVar          = PATIENT,
            LongiModel          = SLOPES,
            LongiResponse       = CD4,
            LongiTimevar        = OBSTIME,
            LongiTimeInter      = DDI,
            LongiGMatrix        = UN,
            TimetoEventVar      = TIME,
            EventVar            = DEATH,
            EventVal            = 1,
            EventModel          = PIECEWISE,
            EventCovariates     = DDI,
            NLMIXEDOptions      = GCONV=0 QPOINTS=5,
            OutputPredictions   = XIBETA XIBETA_ZIBI,
            OutputPredictdata   = XIBETA XIBETA_ZIBI,
            SharedParam         = CURRENT_VALUE,
            AdditionalOptions   = CALCULATEEXECTIME);
            SharedParam         = CURRENT_VALUE);
```

Section 4 provides the complete listing of macro parameters. Section 6 illustrates the use of the macro with several examples.

4. List of macro parameters

4.1.   Identification of input data

Data=<dataset>

Name of the input dataset. Required

DataStructure=standard|vertical

Structure of the input dataset.

> **STANDARD**: input data with separate columns for longitudinal and time-to-event
>
> responses (default).
>
> **VERTICAL**: input data structured in separate rows for longitudinal and time-to-event
>
> responses. See further details in section 6.4.

Where = <condition>

Condition to select a subset of rows.

SubjectVar = <variable>

Name of the variable that identifies the subject. Required.

4.2.   Identification of output

OutputParameters =<library.dataset>

Dataset to store the estimated model parameters. By default, the estimated parameters are saved in work. jmoutputparameters. See further details on the structure of this dataset in section 7.3.

Outputpredictions =<list of model terms>

Requests the macro the creation of datasets with predictions. Available terms are: XiBeta, XiBeta_Zibi, mi, log_hi and cum_hi. It is possible to specify a transformation of these such as exp(-cum_hi). When the term depends not only on fixed effects but also on random effects, the empirical Bayes estimates of these are used.

OutputpredictData =<list of dataset names>

List of names of the (libraries and) datasets used to save the predictions.

ListingFile = 'file name'

File to send the output by the ods listing statement. Note that If UNTOUCHEDLISTING is added to the AdditionalOptions macro parameter, all macro ods listing statements are disabled.

4.3. Set-up of the longitudinal model

LongiType = NORMAL |BINARY | BINOMIAL |POISSON

Distribution of the longitudinal response $y_i(t)$. Normal is used by default.

LongiLink = IDENTITY |LOGIT | LOG | PROBIT

Link function. The canonical link is used by default.

LongiVar = <var name>

Name of the variable with the longitudinal response. Required.

LongiEventCat = <category>

For binary longitudinal responses only, category of interest. Default value is 1.

LongiTrialsVar = <var name>

For binomial longitudinal responses only, name of the variable with number of trials.

LongiOffset = <var name>

Name of the variable to be included in the linear model with a coefficient fixed to 1.

LongiTimeVar = <var name>

Name of the variable with the time at which the longitudinal response has been observed. Required.

LongiTimeModel=LINEAR|SPLINE|NATURALCUBIC|BSLINE

Approach to fitt the trajectories of the longitudinal response versus time. Required.

LINEAR: Fits a random intercepts and slopes model. This is the option used by default.

SPLINE: Fits a random splines model.

NATURALCUBIC: Fits a random natural cubic splines model.

BSPLINE: Fits a random B-splines model.

LongiLowerKnot = <time> LongiUpperKnot = <time>

Lower and upper boundary knots for the modeling of the longitudinal response using B-splines or natural cubic splines.

LongiNKnots = <positive integer>

Number of internal knots for the modeling of the longitudinal response using splines, B-splines or natural cubic splines. No internal knots are assumed by default.

LongiKnot1 = <time>,..., LongiKnot9 = <time>X

Internal knots for the modeling of the longitudinal response using splines. If not given by user an equally spaced quantile partition is used.

LongiCovariates = <var list>

List of baseline covariates of the longitudinal response model.

LongiTimeInter = <var name>

Name of the variable that interacts with the longitudinal trajectory.

LongiGMatrix=UN|VC|ANTE(1)

Structure of the G matrix.

VC: random effects are uncorrelated. This is the option used by default.

UN: unstructured covariance structure.

ANTE(1): antedependence covariance structure.

LongiGMatrixParam =VARCOVAR | CHOL |VARCORR

Parameterization of the G matrix.

> **VARCOVAR**: parameters for variances and covariances (default).
>
> **CHOL**: parameters for the Cholesky decomposition.
>
> **VARCORR**: parameters for variances and correlations. Required option when LongiGMatrix= ANTE(1) .

LongiSmallLogSD = <value>

Value to use as initial value for the log-transformed standard deviation of random effects where the disjoint linear model has estimated 0. Default value is -10.

LongiModelOptions =<options>

Options in the MIXED (or GLIMMIX) statement of the MIXED (or GLIMMIX) procedure used to estimate the initial parameters. For normal responses, by default, METHOD=ML is used.

LongiGLINLOptions =<options>

For non-normal responses, options in the NLOPTIONS statement of the GLIMMIX procedure used to estimate the initial parameters.

4.4.   Set-up of the time-to-event model

EventTimeVar = <var name>

Name of the variable with the time at which the event or the censoring occurs. Required.

EventVar = <var name>

Name of the variable with the indicator of event or censoring. Required.

EventVal = <category>

Category that indicates event of interest. All other categories are assumed censored observations. Required.

EventModel = EXPONENTIAL | WEIBULL | PIECEWISE | GSPLINE | GBSPLINE | GNATURALCUBIC | WSPLINE | WBSPLINE | | RPSPLINE | RPBSPLINE | RPNATURALCUBIC

Distribution of the baseline hazard in the time-to-event model. Required.

EventTimeStart = <var name>

For data in vertical structure, name of the variable with the start of the interval.

EventStrata = <var name>

For stratified models, name of the variable to identify the strata.

EventNKnots = <positive integer>

Number of internal knots for the modeling of the time-to-event responses using a piecewise exponential model or splines. Default value is 6.

EventKnot1 = <time>, ...,EventKnot9 = <time>X

Internal knots for the modeling of the time-to-event responses using a piecewise exponential model or splines. If not given by user, an equally spaced quantile partition is used.

EventCovariates = <var list>

List of baseline covariates of the time-to-event model.

4.5.   Set-up of the joint model

InitialParameters =disjoint | <dataset>

Option to estimate initial parameters.

> **DISJOINT**: uses parameters obtained from the disjoint longitudinal and time-to-event submodels (default);
>
> **<dataset>**: uses a dataset where a set of initial parameters has been stored, e.g. obtained from a previous execution of the macro.

SharedParam = CURRENT_VALUE | SLOPE | CUMULATIVE | COEFFICIENTS | BASELINE_VALUE | CHANGE_FROM_BASELINE

Parameterization to associate the longitudinal model with the time-to-event model. Multiple choices are possible. The current-value parameterization is used by default.

SharedLagging = <positive real>

Lagging in the association between the longitudinal and the time-to-event response. Default value is 0.

SharedParamInt = <var list>

List of variables that interact with the effect of the longitudinal response on the event risk model.

SharedCoefficients = bi0 |bi1 | bi2 | bi3 | ...

For the random-effects coefficients parameterization, random effects used as shared parameters.

Required if SharedParam = coefficients.

SharedLongiTerm= MEAN | LINEAR_PREDICTOR | COEFFICIENTS

Option to link the longitudinal and the time-to-event models. See section 2.3.8 for further details.

> **MEAN**: a linear relationship between the log-hazard function and the expected response of the longitudinal response $m_i(t)$ [or $m'_i(t)$]. This is the option used by default if SharedParam = CURRENT_VALUE [or SLOPE]. Note: this option is not available if SharedParam = CUMULATIVE.
>
> **LINEAR_PREDICTOR**: the log-hazard function linearly related to the linear predictor $g\{m_i(t)\}$.
>
> **COEFFICIENTS:** a linear relationship between the log-hazard function and the individual coefficients. This is the only possible option if SharedParam = COEFFICIENTS.

NLMIXEDOptions = <options>

Options in the NLMIXED statement of the NLMIXED procedure used to fit the joint model.

4.6.   Other parameters

AdditionalOptions = <option>

Requests the macro additional options. The following options can be requested:

> **CALCULATEEXECTIME**: calculates and displays the macro execution time.
>
> **COMPETING**: informs the macro that the variable given in the EventStrata parameter
> defines competing events. With this option, the macro expects a vertical data
> structure where rows have been included for each of the competing time-to-event
> responses.
>
> **EVENTTIMES**: uses only actual (non-censored) events to calculate the time partition
> for the time-to-event model. Applicable for the piecewise exponential model and
> the time-to-event models using splines.
>
> **FITSTATISTICS**: creates a dataset with the fit statistics. The dataset is placed in the
> same location and with the same name as the output parameters dataset,
> adding the suffix '_fit'.
>
> **ONLYEVENT**: fits only the time-to-event model.
>
> **ONLYLONGI**: fits only the longitudinal model.
>
> **NOINT**: removes the intercept from the longitudinal model.
>
> **NOMACROWARN**: disables all macro warnings.
>
> **SKIPMACROHEADER**: skips the header produced by the macro before the execution of
> the NLMIXED procedure.
>
> **UNTOUCHEDLISTING**: disables all macro ods listing statements.

SmallTimeDif = <time>

Amount of time small enough to adequately estimate the first derivatives using the central difference
approximation. 0.001 is used by default.

NLMixedExpression= NO | YES

In case the user wants to add an expression inside the NLMIXED procedure this has to be set to YES.
User needs to define a macro called NLMixedExpression before the call to %JM including the new
expression inside the macro.

For example in the example below defines Malebaseline before the call to%JM so later they can request predictions of XiBetaSq.

```
%macro NLMixedExpression;
MaleBaseline=B0+bio+SEX;
%mend;
%JM(…..,Outputpredictions= Malebaseline, OutputpredictData = MalebaselineData);
```

NLMixedExpression= NO | YES

In case the user wants to add an expression inside the NLMIXED procedure this has to be set to YES.

Random_Effects_in_Input_data= NO | YES

This allows to call the %JM macro without the RANDOM statement in Proc NLMixed.. This assumes that fixed values for the random effects will be provided in the input data. So the use has added variables such as bi0, bi1, bi2, etc.

5. JM package versus %JM macro: a comparative overview

The table below provides a comparative overview of the functionalities of the %JM macro versus the JM
R package [5, 6].

Table 4. JM R package versus %JM SAS Macro

	JM R package V1.2-0	%JM SAS macro V 1.04
Longitudinal Model		
Trajectories over time: linear, polynomials, splines, B-splines, natural cubic splines	X	X
Baseline Covariates	X	X
Time-interaction terms	X	X
Time-to-Event Model		
PH Weibull	X	X
PH Piecewise-exponential	X	X
PH Unspecified baseline	X	
PH Generalization of Gompertz based on splines	X	X
PH Generalization of Weibull based on splines		X
PH Royston-Parmar model	X [5]	X
AFT Weibull	X	
Exogenous time-dependent covariates	X	X
Stratification Factors	X	X
Parameterizations to link both models		
Time-dependant (trajectory Model)	X	X
Slope-dependant	X	X
Cumulative-dependant	X	X
Lagging	X	X
Random-effects coefficients	X	X
Interaction effects	X	X
Approximation of the survival function		
7-points Gauss-Kronrod Rule	X	
15-points Gauss-Kronrod Rule	X	X
Approximation of the integral over the random effects		
Gauss-Hermitage	X	
Adaptive Gauss-Hermitage		X
Pseudo-adaptive Gauss-Hermitage	X	
Laplace Approximation	X	X
Optimization Techniques		
EM algorithm	X	
Quasi-Newton algorithm	X	X
Other methods available in NLMIXED		X

---

[5] For method = "ch-Laplace", the R package uses an additive model on the log cumulative hazard scale with B-splines, instead of natural cubic splines as proposed by Royston and Parmar.

6. Examples

6.1. Example 1: a simple current-value shared-parameter model

In this section we fit the joint model over the aids dataset [17] illustrated in page 59 of Rizopoulos 2012 [1]. The CD4 is fitted using a random-slopes model with a separated slope fitted for subjects with DDI treatment. An unstructured covariance was assumed for the G matrix so an extra parameter captures the covariance between b_{i0} and b_{i1}.

$$m_i(t) = \beta_0 + \beta_1 t + \beta_2 t\, DDI_i + b_{i0} + b_{i1} t \qquad [6.1.1]$$

The time-to-death was fitted using a piecewise exponential model using a partition of 7 intervals (6 knots), used by default by the macro, with DDI as baseline covariate. Both models were joined using a (trajectory) current-value shared-parameter as given by

$$h_i(t) = h_0(t) \exp\{\gamma_1 DDI_i + \alpha\, m_i(t)\} \qquad [6.1.2]$$

For such a model the SAS code is given by

```
%JM(        Data              = AIDS,
            SubjectVar        = PATIENT,
            LongiTimeModel    = LINEAR,
            LongiVar          = CD4,
            LongiTimevar      = OBSTIME,
            LongiTimeInter    = DDI,
            LongiGMatrix      = UN,
            TimetoEventVar    = TIME,
            EventVar          = DEATH,
            EventVal          = 1,
            EventModel        = PIECEWISE,
            EventCovariates   = DDI,
            NLMIXEDOptions    = GCONV=0 QPOINTS=5,
            OutputPredictions = XIBETA XIBETA_ZIBI,
            OutputPredictdata = XIBETA XIBETA_ZIBI,
            SharedParam       = CURRENT_VALUE,
            AdditionalOptions = CALCULATEEXECTIME);
```

For this model, the %JM macro provides the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Normal. Link function: identity.
Longitudinal Process: Random Intercepts and Slopes Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: current_value.
Event Process:      467 subjects,       188 events.

                        The NLMIXED Procedure

                        Specifications

Data Set                              WORK._JMDATA
Dependent Variable                    JMResponse
Distribution for Dependent Variable   General
Random Effects                        bi0 bi1
Distribution for Random Effects       Normal
Subject Variable                      patient
Optimization Technique                Dual Quasi-Newton
Integration Method                    Adaptive Gaussian
                                      Quadrature


                  Dimensions

Observations Used               1872
Observations Not Used              0
Total Observations              1872
Subjects                         467
Max Obs Per Subject                6
Parameters                        16
Quadrature Points                  5


               Iteration History

Iter   Calls   NegLogLike      Diff    MaxGrad      Slope

   1       4   4373.30009   15.71245  103.0421   -3030.82
   .       .        .            .         .          .
  41      73   4328.38921   4.64E-11  1.279E-6   -889E-13


   NOTE: FCONV convergence criterion satisfied.

             Fit Statistics

   -2 Log Likelihood              8656.8
   AIC (smaller is better)        8688.8
   AICC (smaller is better)       8689.1
   BIC (smaller is better)        8755.1

                                 Additional Estimates

                                            Standard
Label                              Estimate    Error   DF   t Value   Pr > |t|

Longitudinal Process: Intercept      7.2048   0.2218  465    32.49    <.0001
Longitudinal Process: Slope         -0.1921   0.02173 465    -8.84    <.0001
Longitudinal Process: DDI by slope   0.01140  0.03026 465     0.38    0.7067
Longitudinal Process: Log SD Residual          0.5524   0.02749 465    20.09    <.0001
Longitudinal Process: Log SD Random Intercept  1.5223   0.03577 465    42.56    <.0001
Longitudinal Process: Log SD Random Slope     -1.7003   0.09677 465   -17.57    <.0001
Longitudinal Process: Covariance Intercept by Slope  -0.03784  0.07188 465  -0.53  0.5988
Event Process: Log Hazard Period 1  -2.5160   0.1915  465   -13.14    <.0001
Event Process: Log Hazard Period 2  -2.2994   0.1788  465   -12.86    <.0001
Event Process: Log Hazard Period 3  -1.9222   0.2402  465    -8.00    <.0001
Event Process: Log Hazard Period 4  -2.5008   0.3423  465    -7.31    <.0001
Event Process: Log Hazard Period 5  -2.4188   0.3156  465    -7.66    <.0001
Event Process: Log Hazard Period 6  -2.4318   0.4018  465    -6.05    <.0001
Event Process: Log Hazard Period 7  -2.4121   0.5313  465    -4.54    <.0001
Event Process: DDI                   0.3349   0.1565  465     2.14    0.0329
Association (current-value)         -0.2892   0.03603 465    -8.03    <.0001

Execution Time (hours:minutes:seconds):  0:11:36.
```

The parameter OutputPredictions has been used to request the macro to save both the population predictions and the empirical Bayes estimates of the subject predictions. The code below produces the two plots shown in figure 1, the first plot includes both all individual observations and the estimated longitudinal response for each treatment whereas the second plot gives the individual observations and the empirical Bayes estimates of the longitudinal response for patient 10.

```
proc sgplot data= XIBETA;
  series   x=OBSTIME y=PRED  /legendlabel="Predictions"   group=DDI;
  scatter  x=OBSTIME y=CD4  /legendlabel="Observations" group=DDI;
  where JMDIST ="LONGI" ;

proc sgplot data= XIBETA_ZIBI;
  series   x=obstime y=Pred  /legendlabel="Predictions";
  scatter  x=obstime y=CD4  /legendlabel="Observations";
  yaxis values=(0 to 25 by 5);
  where JMDIST ="LONGI" and PATIENT=10;
```

Figure 1. CD4 Cell Count versus time by treatment
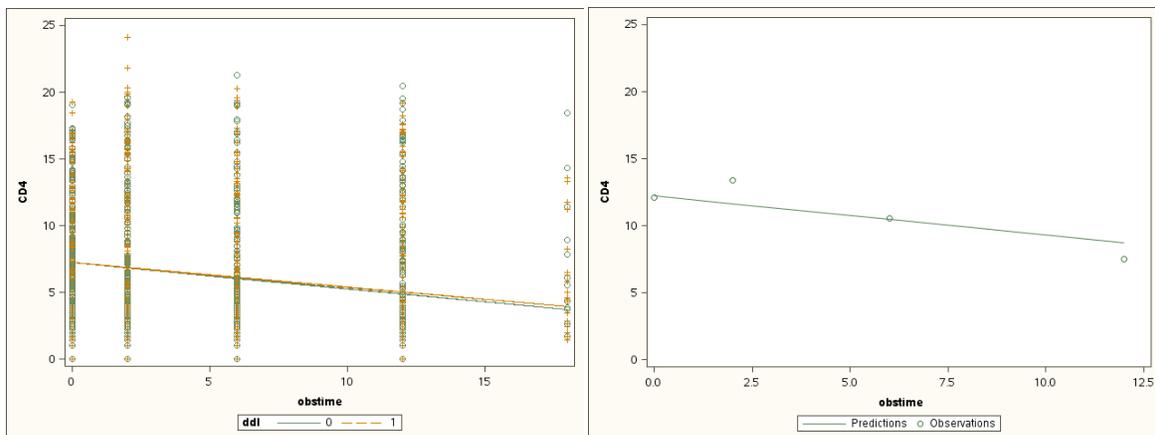
6.2.   Example 2: a model with non-linear trajectories and interaction effects

There are occasions where the effect of the longitudinal response on the risk of event is not the same for all subjects. To illustrate how the %JM macro can be used to fit a model with interactions in the association between the longitudinal model and the time-to-event model, we use the example in primary biliary cirrhosis [18] illustrated in page 98 in Rizopoulos 2012 [1]. The log-transformed serum bilirubin was analyzed with a linear mixed model with a quadratic evolution over time with a different evolution fitted for patients treated with penicillin with respect to the ones treated with placebo:

$$m_i(t) = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t + (\beta_2 + b_{i2})t^2 + \beta_3 D\_PNC_i + \beta_4 t\, D\_PNC_i + \beta_5 t^2 D\_PNC_i \qquad [6.2.1]$$

An unstructured covariance was assumed for the G matrix what means that three extra parameters captured the covariance between b_{i0}, b_{i1} and b_{i2}. The time-to-death was fitted using a piece-wise exponential model with a 7-interval equally-spaced quantile partition, and drug (D_PNC) and hepatomegaly (HMG) as baseline covariates. Both models were joined using two (trajectory) current-value shared-parameters as given below:

$$h_i(t) = h_0(t) \exp[\gamma_1 \, D\_PNC_i + \gamma_2 \, HMG_i + \alpha_1 \, m_i(t) + \alpha_2 \, \{m_i(t) \times HMG_i\}] \qquad [6.2.2]$$

In order to fit this model with %JM, we use the following call to the macro:

```
%JM(   Data              = PBC,
       SubjectVar        = ID,
       LongiResponse     = LOGSERBILIR,
       LongiTimeModel    = SPLINE,
       LongiTimevar      = YEAR,
       LongiTimeInter    = D_PNC,
       LongiDegree       = 2,
       LongiCovariates   = D_PNC,
       LongiGMatrix      = UN,
       TimetoEventVar    = YEARS,
       EventVar          = NOALIVE,
       EventVal          = 1,
       EventModel        = PIECEWISE,
       EventCovariates   = D_PNC HMG,
       SharedParam       = CURRENT_VALUE,
       SharedParamInt    = HMG,
       NLMIXEDOptions    = GCONV=0 QPOINTS=3,
       AdditionalOptions = CALCULATEEXECTIME);
```

For this model, the %JM macro provides the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Normal. Link function: identity.
Longitudinal Process: Random Splines Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: current_value.
Event Process:     312 subjects,     169 events.


              The NLMIXED Procedure

              Specifications

Data Set                              WORK._JMDATA
Dependent Variable                    JMResponse
Distribution for Dependent Variable   General
Random Effects                        bi0 bi1 bi2
Distribution for Random Effects       Normal
Subject Variable                      id
Optimization Technique                Dual Quasi-Newton
Integration Method                    Adaptive Gaussian
```

```
                         Quadrature


                   Dimensions

           Observations Used          2257
           Observations Not Used      1945
           Total Observations         4202
           Subjects                    312
           Max Obs Per Subject          17
           Parameters                   24
           Quadrature Points             3


                 Iteration History

Iter    Calls    NegLogLike       Diff     MaxGrad      Slope

  1        3    1941.05235     50.96588    80.60698    -1536.4
  .        .             .            .           .          .
 67      130    1857.68059     2.27E-12    0.000097    -6.62E-6


      NOTE: FCONV convergence criterion satisfied.


                 Fit Statistics

      -2 Log Likelihood              3715.4
      AIC (smaller is better)        3763.4
      AICC (smaller is better)       3763.9
      BIC (smaller is better)        3853.2


                         The NLMIXED Procedure

                         Additional Estimates

                                                   Standard
Label                                   Estimate     Error    DF   t Value   Pr > |t|   Alpha

Longitudinal Process: Intercept           0.5858    0.08246   309     7.10    <.0001     0.05
Longitudinal Process: Time Parameter 1    0.1678    0.03187   309     5.26    <.0001     0.05
Longitudinal Process: Time Parameter 2    0.003109  0.003580  309     0.87    0.3859     0.05
Longitudinal Process: D_PNC by Time Parameter 1  0.000267  0.04375  309  0.01  0.9951   0.05
Longitudinal Process: D_PNC by Time Parameter 2  -0.00201  0.004442  309  -0.45  0.6515  0.05
Longitudinal Process: D_PNC              -0.1432    0.1159    309    -1.24    0.2176     0.05
Longitudinal Process: Log SD Residual    -1.1948    0.02054   309   -58.16   <.0001     0.05
Longitudinal Process: Log SD Random Intercept  -0.00328  0.04244  309  -0.08  0.9384   0.05
Longitudinal Process: Log SD Random Time Parameter 1  -1.1748  0.06495  309  -18.09  <.0001  0.05
Longitudinal Process: Log SD Random Time Parameter 2  -3.6904  0.09679  309  -38.13  <.0001  0.05
Longitudinal Process: Covariance Time Param 0 by Param 1  0.05991  0.02527  309  2.37  0.0184  0.05
Longitudinal Process: Covariance Time Param 0 by Param 2  0.000733  0.003111  309  0.24  0.8138  0.05
Longitudinal Process: Covariance Time Param 1 by Param 2  -0.00660  0.001184  309  -5.58  <.0001  0.05
Event Process: Log Hazard Period 1       -4.7944    0.3310    309   -14.48   <.0001     0.05
Event Process: Log Hazard Period 2       -4.4369    0.3334    309   -13.31   <.0001     0.05
Event Process: Log Hazard Period 3       -4.4403    0.3579    309   -12.41   <.0001     0.05
Event Process: Log Hazard Period 4       -4.2377    0.3825    309   -11.08   <.0001     0.05
Event Process: Log Hazard Period 5       -4.2445    0.3825    309   -11.10   <.0001     0.05
Event Process: Log Hazard Period 6       -4.0862    0.4060    309   -10.06   <.0001     0.05
Event Process: Log Hazard Period 7       -4.5847    0.5167    309    -8.87   <.0001     0.05
Event Process: D_PNC                      0.09389   0.1678    309     0.56    0.5763     0.05
Event Process: HMG                        0.2358    0.4126    309     0.57    0.5681     0.05
Association (current-value)               1.1594    0.1290    309     8.99    <.0001     0.05
Association by HMG (current-value)        0.2328    0.1856    309     1.25    0.2107     0.05

Execution Time (hours:minutes:seconds):  0:44:02.
```

6.3. Example 3: a model with a binary longitudinal response

In this section we fitted a joint model similar to the one shown in section 6.1 but using the binary longitudinal response π_i defined as subject i having the CD4 cell count higher or equal to 30 cells per cubic millimeter. This response was fitted using a random-slopes model with a separated slope fitted for

subjects with DDI treatment. An unstructured covariance was assumed for the G matrix that means that an extra parameter captures the covariance between b_{i0} and b_{i1}.

$$\log\left\{\frac{\pi_i(t)}{1-\pi_i(t)}\right\} = \beta_0 + \beta_1 t + \beta_2 t \, DDI_i + b_{i0} + b_{i1} t \qquad [6.3.1]$$

The time-to-death was fitted using a piecewise exponential model using a partition of 7 intervals (6 knots), used by default by the macro, with DDI as baseline covariate. Both models were joined using a (trajectory) current-value shared-parameter as given by

$$h_i(t) = h_0(t) \exp\{\gamma_1 \, DDI_i + \alpha \, \pi_i(t)\} \qquad [6.3.2]$$

In order to fit the joint model we used 9 quadrature points since the Gauss-Hermite technique is known to require a larger number of quadrature points in the random-effects logistic regression model (19). For such a model the SAS code is given by

```
%JM(       Data              = AIDS,
           SubjectVar        = PATIENT,
           LongiType         = BINARY,
           LongiTimeModel    = LINEAR,
           LongiVar          = CD4_A30,
           LongiTimevar      = OBSTIME,
           LongiTimeInter    = DDI,
           LongiGMatrix      = UN,
           LongiModelOptions = METHOD=QUAD,
           TimetoEventVar    = TIME,
           EventVar          = DEATH,
           EventVal          = 1,
           EventModel        = PIECEWISE,
           EventCovariates   = DDI,
           NLMIXEDOptions    = GCONV=0 QPOINTS=9,
           SharedParam       = CURRENT_VALUE,
           AdditionalOptions = CALCULATEEXECTIME);
```

Unfortunately the execution stopped abruptly showing the following message when the NLMIXED procedure was attempting fitting the joint model:

ERROR: No valid parameter points were found.

This error is often produced when the NLMIXED procedure is not given good set of initial values for the parameters. For non-normal data, the %JM macro uses first the GLIMMIX procedure to fit the longitudinal model, in this case a random-effects logistic regression, and takes the estimates as initial parameters for the joint model. By default, the GLIMMIX procedure uses pseudo-likelihood (PL) estimation methods based on linearization. The PL methods are known to produce biased estimates for binary data when the number of observations per subject is small [20]. In order to provide better initial parameters to the NLMIXED procedure, we should set the LongiModelOptions macro parameter to METHOD=QUAD to force the use of the adaptive Gauss-Hermite approximation also in the initial model fitted to estimate the initial parameters of the longitudinal response. Once this new parameter was added to the macro call, the joint model was fitted with no errors leading to the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Binary. Link function: logit.
Longitudinal Process: Random Intercepts and Slopes Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: current_value.
Event Process:      467 subjects,      188 events.

                   The NLMIXED Procedure

                   Specifications

Data Set                            WORK._JMDATA
Dependent Variable                  JMResponse
Distribution for Dependent Variable General
Random Effects                      bi0 bi1
Distribution for Random Effects     Normal
Subject Variable                    patient
Optimization Technique              Dual Quasi-Newton
Integration Method                  Adaptive Gaussian
                                    Quadrature


                      Dimensions

        Observations Used            1872
        Observations Not Used           0
        Total Observations           1872
        Subjects                      467
        Max Obs Per Subject             6
        Parameters                     15
        Quadrature Points               9


                     Fit Statistics

        -2 Log Likelihood            3014.9
        AIC (smaller is better)      3044.9
        AICC (smaller is better)     3045.2
        BIC (smaller is better)      3107.1


                             Additional Estimates

                                                    Standard
Label                                   Estimate     Error    DF   t Value   Pr > |t|

Longitudinal Process: Intercept          0.5573     0.2515    465     2.22     0.0272
Longitudinal Process: Slope             -0.2425     0.05421   465    -4.47    <.0001
Longitudinal Process: DDI by slope      -0.00133    0.04735   465    -0.03     0.9777
Longitudinal Process: Log SD Random Intercept   1.4275   0.1120   465   12.74   <.0001
Longitudinal Process: Log SD Random Slope      -2.3821   1.2436   465   -1.92    0.0560
Longitudinal Process: Covariance Intercept by Slope   0.2404   0.3006   465   0.80   0.4244
Event Process: Log Hazard Period 1      -3.1200     0.1652    465   -18.88    <.0001
Event Process: Log Hazard Period 2      -2.8634     0.1584    465   -18.08    <.0001
```

```
Event Process: Log Hazard Period 3          -2.4507    0.2262   465  -10.83   <.0001
Event Process: Log Hazard Period 4          -3.0222    0.3319   465   -9.10   <.0001
Event Process: Log Hazard Period 5          -2.9180    0.3035   465   -9.61   <.0001
Event Process: Log Hazard Period 6          -2.8845    0.3895   465   -7.41   <.0001
Event Process: Log Hazard Period 7          -2.7927    0.5100   465   -5.48   <.0001
Event Process: DDI                           0.2885    0.1513   465    1.91   0.0571
Association (current-value)                 -2.1753    0.3630   465   -5.99   <.0001

Execution Time (hours:minutes:seconds):  0:35:11.
```

6.4. Example 4: improving the approximation of the survival function with vertical input data

In this section we will repeat the same joint model fitted in section 6.1 using an input data set that has been structured vertically (as shown in Table 6). As illustrated in section 2.4.1, the Kronrod rule will then be applied to each interval leading to both a more accurate approximation of the survival function and an increased execution time.

For such a model the SAS code is given by

```
%JM(        Data            = AIDSVERTICAL,

            SubjectVar      = PATIENT,

            DataStructure   = VERTICAL,

            LongiTimeModel  = LINEAR,

            LongiVar        = CD4,

            LongiTimevar    = OBSTIME,

            LongiTimeInter  = DDI,

            LongiGMatrix    = UN,

            EventTimeVar    = TOTIME,

            EventTimeStart  = FROMTIME,

            EventVar        = DEATH2,

            EventVal        = 1,

            EventModel      = PIECEWISE,

            EventCovariates = DDI,

            NLMIXEDOptions  = GCONV=0 QPOINTS=5,

            SharedParam     = CURRENT_VALUE,

            AdditionalOptions = CALCULATEEXECTIME);
```

For this model, the %JM macro provides the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Normal. Link function: identity.
Longitudinal Process: Random Intercepts and Slopes Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: current_value.
Event Process:    467 subjects,    188 events.

                   The NLMIXED Procedure

                      Specifications

Data Set                            WORK._JMDATA
Dependent Variable                  JMResponse
```

```
Distribution for Dependent Variable        General
Random Effects                             bi0 bi1
Distribution for Random Effects            Normal
Subject Variable                           patient
Optimization Technique                     Dual Quasi-Newton
Integration Method                         Adaptive Gaussian
                                           Quadrature


              Dimensions

Observations Used              3266
Observations Not Used             0
Total Observations             3266
Subjects                        467
Max Obs Per Subject              12
Parameters                       16
Quadrature Points                 5


              Iteration History

Iter    Calls    NegLogLike      Diff     MaxGrad       Slope

   1       4    4373.34496    15.73627   102.4166    -3038.42
   .       .            .           .          .           .
  39      70    4328.59044    3.21E-10   0.000012    -629E-12


      NOTE: FCONV convergence criterion satisfied.


              Fit Statistics

    -2 Log Likelihood            8657.2
    AIC (smaller is better)      8689.2
    AICC (smaller is better)     8689.3
    BIC (smaller is better)      8755.5
```

              Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Longitudinal Process: Intercept	7.2045	0.2218	465	32.49	<.0001
Longitudinal Process: Slope	-0.1920	0.02173	465	-8.84	<.0001
Longitudinal Process: DDI by slope	0.01150	0.03027	465	0.38	0.7041
Longitudinal Process: Log SD Residual	0.5524	0.02749	465	20.09	<.0001
Longitudinal Process: Log SD Random Intercept	1.5224	0.03577	465	42.56	<.0001
Longitudinal Process: Log SD Random Slope	-1.7001	0.09673	465	-17.57	<.0001
Longitudinal Process: Covariance Intercept by Slope	-0.03823	0.07189	465	-0.53	0.5951
Event Process: Log Hazard Period 1	-2.5221	0.1915	465	-13.17	<.0001
Event Process: Log Hazard Period 2	-2.2872	0.1785	465	-12.81	<.0001
Event Process: Log Hazard Period 3	-1.9595	0.2403	465	-8.15	<.0001
Event Process: Log Hazard Period 4	-2.4993	0.3413	465	-7.32	<.0001
Event Process: Log Hazard Period 5	-2.4142	0.3157	465	-7.65	<.0001
Event Process: Log Hazard Period 6	-2.4005	0.4009	465	-5.99	<.0001
Event Process: Log Hazard Period 7	-2.4221	0.5302	465	-4.57	<.0001
Event Process: DDI	0.3333	0.1565	465	2.13	0.0337
Association (current-value)	-0.2888	0.03602	465	-8.02	<.0001

```
Execution Time (hours:minutes:seconds):  0:48:14.
```

6.5. Example 5: a model with two shared-parameters: baseline-value and (trajectory) change-from-baseline

In this section we fit a joint model over the aids dataset very similar to the one show in example 1 . An unstructured covariance was assumed for the G matrix so an extra parameter captures the covariance between b_{i0} and b_{i1}.

$$m_i(t) = \beta_0 + \beta_1 t + \beta_2 t\, DDI_i + b_{i0} + b_{i1} t \qquad\qquad [6.5.1]$$

The time-to-death was fitted using a piecewise exponential model using a partition of 7 intervals (6 knots), used by default by the macro, with DDI as baseline covariate.

Both models will be joined using both a baseline-value shared-parameter and a (trajectory) change-from-baseline shared-parameter as given by

$$h_i(t) = h_0(t) \exp\{\gamma_1\, DDI_i + \alpha_1\, m_i(0) + \alpha_2\, ch\_m_i(t)\} \qquad\qquad [6.5.2]$$

For such a model the SAS code is given by

```
%JM(        Data             = AIDS,
            SubjectVar       = PATIENT,
            LongiTimeModel   = LINEAR,
            LongiVar         = CD4,
            LongiTimevar     = OBSTIME,
            LongiTimeInter   = DDI,
            LongiGMatrix     = UN,
            TimetoEventVar   = TIME,
            EventVar         = DEATH,
            EventVal         = 1,
            EventModel       = PIECEWISE,
            EventCovariates  = DDI,
            NLMIXEDOptions   = GCONV=0 QPOINTS=5,
            OutputPredictions = XIBETA XIBETA_ZIBI,
            OutputPredictdata = XIBETA XIBETA_ZIBI,
            SharedParam      = BASELINE_VALUE CHANGE_FROM_BASELINE,
            AdditionalOptions = CALCULATEEXECTIME);
```

For this model, the %JM macro provides the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Normal. Link function: identity.
Longitudinal Process: Random Intercepts and Slopes Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: baseline_value change_from_baseline.
Event Process:     467 subjects,      188 events.

                     The NLMIXED Procedure

                         Specifications

Data Set                              WORK._JMDATA
Dependent Variable                    JMResponse
Distribution for Dependent Variable   General
Random Effects                        bi0 bi1
Distribution for Random Effects       Normal
Subject Variable                      patient
Optimization Technique                Dual Quasi-Newton
Integration Method                    Adaptive Gaussian
                                      Quadrature


                            Dimensions

          Observations Used              1872
          Observations Not Used             0
          Total Observations             1872
          Subjects                        467
          Max Obs Per Subject               6
          Parameters                       17
          Quadrature Points                 5


                       Iteration History

   Iter    Calls    NegLogLike      Diff     MaxGrad     Slope

      1       4     4377.9315    11.08104   94.65088   -2933.04
     46      85     4321.81572   1.56E-10   2.438E-6   -319E-12


      NOTE: FCONV convergence criterion satisfied.


                       Fit Statistics

      -2 Log Likelihood                 8643.6
      AIC (smaller is better)           8677.6
      AICC (smaller is better)          8678.0
      BIC (smaller is better)           8748.1
```

                          Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
Longitudinal Process: Intercept	7.2244	0.2209	465	32.70	<.0001
Longitudinal Process: Slope	-0.2064	0.02110	465	-9.78	<.0001
Longitudinal Process: DDI by slope	-0.00458	0.02896	465	-0.16	0.8745
Longitudinal Process: Log SD Residual	0.5622	0.02751	465	20.44	<.0001
Longitudinal Process: Log SD Random Intercept	1.5179	0.03580	465	42.40	<.0001
Longitudinal Process: Log SD Random Slope	-1.7154	0.1006	465	-17.04	<.0001
Longitudinal Process: Covariance Intercept by Slope	0.03649	0.07005	465	0.52	0.6027
Event Process: Log Hazard Period 1	-2.7474	0.2271	465	-12.10	<.0001
Event Process: Log Hazard Period 2	-3.1622	0.3544	465	-8.92	<.0001
Event Process: Log Hazard Period 3	-3.0327	0.4443	465	-6.83	<.0001
Event Process: Log Hazard Period 4	-3.6784	0.5283	465	-6.96	<.0001
Event Process: Log Hazard Period 5	-3.7497	0.5475	465	-6.85	<.0001
Event Process: Log Hazard Period 6	-3.8526	0.6299	465	-6.12	<.0001
Event Process: Log Hazard Period 7	-3.7900	0.7403	465	-5.12	<.0001
Event Process: DDI	0.3355	0.1948	465	1.72	0.0856
Association (baseline-value)	-0.3205	0.04224	465	-7.59	<.0001
Association (change-from-baseline)	-0.7029	0.1315	465	-5.35	<.0001

```
Execution Time (hours:minutes:seconds):  0:22:29.
```

6.6.　　Example 6: a model with transformed (trajectory) current-value shared parameter

In this section we fit a joint model over the aids dataset very similar to the one show in example 1 . An unstructured covariance was assumed for the G matrix so an extra parameter captures the covariance between b_{i0} and b_{i1}.

$$m_i(t) = \beta_0 + \beta_1 t + \beta_2 t\, DDI_i + b_{i0} + b_{i1} t \qquad [6.6.1]$$

The time-to-death was fitted using a piecewise exponential model using a partition of 7 intervals (6 knots), used by default by the macro, with DDI as baseline covariate.

Both models will be joined using the following equation:

$$h_i(t) = h_0(t)\, \exp\left\{\gamma_1\, DDI_i + \alpha\, \{m_i(t) - 10\}^2\right\} \qquad [6.6.2]$$

For such a model the SAS code is given by

```
%macro transform_current_value_term(cvt=);
((&cvt.- 10)**2)
%mend;
%JM(        Data             = AIDS,
            SubjectVar       = PATIENT,
            LongiTimeModel   = LINEAR,
            LongiVar         = CD4,
            LongiTimevar     = OBSTIME,
            LongiTimeInter   = DDI,
            LongiGMatrix     = UN,
            TimetoEventVar   = TIME,
            EventVar         = DEATH,
            EventVal         = 1,
            EventModel       = PIECEWISE,
            EventCovariates  = DDI,
            NLMIXEDOptions   = GCONV=0 QPOINTS=5,
            OutputPredictions  = XIBETA XIBETA_ZIBI,
            OutputPredictdata  = XIBETA XIBETA_ZIBI,
            SharedParam      = CURRENT_VALUE,
            AdditionalOptions  = CALCULATEEXECTIME,
            Transform_Current_Value_Term = YES);
```

For this model, the %JM macro provides the following output:

```
Joint Model Summary:
Longitudinal Response Distribution: Normal. Link function: identity.
Longitudinal Process: Random Intercepts and Slopes Model
Event Process: Piecewise-constant baseline risk function.
Parameterization: current_value.
Event Process:     467 subjects,      188 events.
The current value was included in the log-risk function after the following transformation: f(x)=((x- 10)**2).
                        The NLMIXED Procedure

                        Specifications

Data Set                              WORK._JMDATA
Dependent Variable                    JMResponse
Distribution for Dependent Variable   General
Random Effects                        bi0 bi1
Distribution for Random Effects       Normal
Subject Variable                      patient
Optimization Technique                Dual Quasi-Newton
Integration Method                    Adaptive Gaussian
                                      Quadrature

                        Dimensions

            Observations Used              1872
            Observations Not Used             0
            Total Observations             1872
            Subjects                        467
            Max Obs Per Subject               6
            Parameters                       16
            Quadrature Points                 5


                    Iteration History

   Iter    Calls    NegLogLike      Diff     MaxGrad     Slope

      1        5     4376.4675   12.54504   1007.156    -101294
     59      131    4332.54132   2.06E-10   0.000891   -0.02241


     NOTE: FCONV convergence criterion satisfied.


                    Fit Statistics

        -2 Log Likelihood             8665.1
        AIC (smaller is better)       8697.1
        AICC (smaller is better)      8697.4
        BIC (smaller is better)       8763.4


                    Additional Estimates
                                         Standard
Label                                    Estimate     Error    DF   t Value  . . . .

Longitudinal Process: Intercept            7.2049    0.2217   465     32.50
Longitudinal Process: Slope               -0.2076   0.02182   465     -9.51
Longitudinal Process: DDI by slope       0.006229   0.03046   465      0.20
Longitudinal Process: Log SD Residual      0.5537   0.02724   465     20.32
Longitudinal Process: Log SD Random Intercept  1.5225  0.03573  465   42.61
Longitudinal Process: Log SD Random Slope -1.6613   0.09209   465    -18.04
Longitudinal Process: Covariance Intercept by Slope  0.01118  0.07242  465  0.15
Event Process: Log Hazard Period 1        -5.3032    0.2574   465    -20.60
Event Process: Log Hazard Period 2        -5.1594    0.2773   465    -18.61
Event Process: Log Hazard Period 3        -4.8074    0.3256   465    -14.76
Event Process: Log Hazard Period 4        -5.3913    0.4078   465    -13.22
Event Process: Log Hazard Period 5        -5.3416    0.3992   465    -13.38
Event Process: Log Hazard Period 6        -5.3775    0.4810   465    -11.18
Event Process: Log Hazard Period 7        -5.3059    0.6149   465     -8.63
Event Process: DDI                         0.3282    0.1774   465      1.85
Association (current-value)               0.03204  0.003429   465      9.35

Execution Time (hours:minutes:seconds):  0:24:56.
```

6.7.   Example 7: a model with a Poisson longitudinal response

In this section we fitted a joint model similar to the one shown in section 6.1 but using the Poisson longitudinal response π_i.

```
*In this dataset we create cd4_as_Poisson assuming it is more or less
a Poisson variable with offset called offset_var;

data mydata.aidspoisson;
set mydata.aids;
cd4_as_Poisson=int(cd4/3);
offset_var=max(int(rand('NORMAL',20,2)),cd4);

run;
```

This response was fitted using a random-slopes model with a separated slope fitted for subjects with DDI treatment. An unstructured covariance was assumed for the G matrix that means that an extra parameter captures the covariance between b_{i0} and b_{i1}.

$$\log\{m_i(t)\} = \beta_0 + \beta_1 t + \beta_2 t\, DDI_i + b_{i0} + b_{i1} t \qquad [6.7.1]$$

The time-to-death was fitted using a piecewise exponential model using a partition of 7 intervals (6 knots), used by default by the macro, with DDI as baseline covariate. Both models were joined using a (trajectory) current-value shared-parameter as given by

$$h_i(t) = h_0(t)\, \exp\{\gamma_1\, DDI_i + \alpha\, m_i(t)\} \qquad [6.7.2]$$

In order to fit the joint model we used 9 quadrature points since the Gauss-Hermite technique is known to require a larger number of quadrature points in generalized random-effects model (19). For such a model the SAS code is given by

```
%JM(       Data                = AIDSPOISSON,
           SubjectVar          = PATIENT,
           LongiType           = POISSON,
           LongiTimeModel      = LINEAR,
           LongiVar            = CD4_AS_POISSON,
           LongiTimevar        = OBSTIME,
           LongiTimeInter      = DDI,
           LongiGMatrix        = UN,
           LongiModelOptions   = METHOD=QUAD(QPOINTS=3),
           LongiOffset         = OFFSET_VAR,
```

```
TimetoEventVar        = TIME,
EventVar              = DEATH,
EventVal              = 1,
EventModel            = PIECEWISE,
EventCovariates       = DDI,
NLMIXEDOptions        = GCONV=0 QPOINTS=9,
SharedParam           = CURRENT_VALUE,
AdditionalOptions     = CALCULATEEXECTIME);
```

6.8. Example 8: calculating dynamic predictions

In this section we will show the code necessary to perform dynamic predictions. First we are going to fit the following model:

```
%JM(Data = mydata.aids,
    SubjectVar = patient,
     LongiType  = normal,
    LongiTimeModel = linear,
    LonqiVar = cd4,
    LongiTimevar = obstime,
    LongiTimeInter = ddi,
    LongiGMatrix = un,
    EventTimeVar = time,
    EventVar = death,
    EventVal = 1,
    EventModel = rpnaturalcubic,
    EventCovariates = ddi,
    InitialParameters = disjoint,
    NLMIXEDOptions     = qpoints=3,
    OutputParameters = Example_params,
    SharedParam = current_value,
    AdditionalOptions = calculateexectime);
```

For the three following subjects we want to estimate its "personalized" predicted survival given the knowledge accumulated for these subjects and the fitted model above.

Patient	ddi	Time	Death	obstime	cd4
Subject 1	1	6.1	0	0	5
				6	7
Subject 2	0	12.5	0	0	3
				6	6
				12	9
Subject 3	0	12.5	0	0	7
				6	5
				12	2

If we know the subject is alive at time t we want to estimate the probability of survival at times s (s>t) given our knowledge that the subject is alive at time t and our knowledge on all longitudinal responses for this subject. Rizopoulos provides on page 173 of his book [3] a first-order estimate for this probability:

$$\pi(u|t) = \Pr\{T_i^* \geq u \mid T_i^* > t;\, Y_i(t);\, \theta\} \cong \frac{S_i\{u \mid M_i(u, \hat{b}_i^{(t)}, \hat{\theta});\hat{\theta}\}}{S_i\{t \mid M_i(t, \hat{b}_i^{(t)}, \hat{\theta});\hat{\theta}\}} \tag{6.8.1}$$

The code that produces dynamic predictions for these subjects is provided in the **Example8RM.sas** program. The resulting dynamic predictions are:
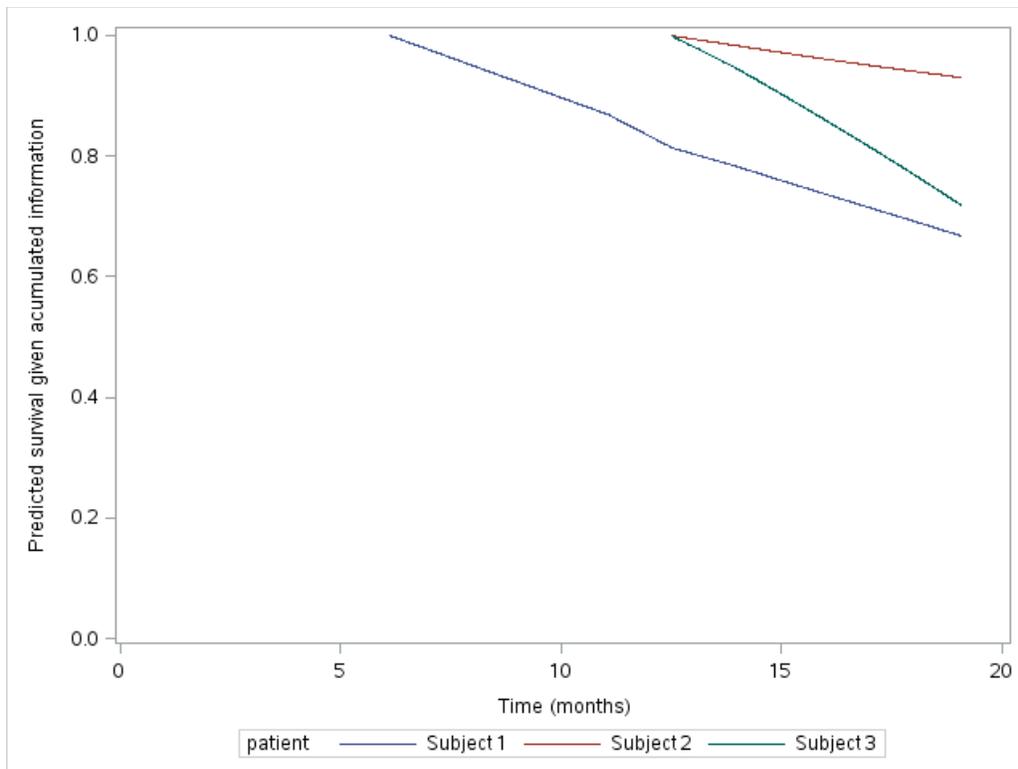
7. Input and output datasets

7.1. Standard input dataset

The standard structure of the input dataset is illustrated in table 4.

Table 5. Standard Input Dataset (aids.sas7bdat)

Id	Obstime	CD4	Time	Death
1	0	10.7	16.97	0
1	6	8.4	16.97	0
1	12	9.4	16.97	0
2	0	6.3	19.00	0
2	6	8.1	19.00	0
2	12	4.6	19.00	0
2	18	5.0	19.00	0

The input dataset needs to have, as a minimum, the following five variables:

1. Subject identification: SubjectVar=<Variable>
2. Time of observation of the longitudinal response: LongiTimevar=<Variable>
3. Longitudinal response value: LongiResponseVar=<Variable>
4. Time To Event: TimetoEventVar=<Variable>
5. Event Indicator: EventVar=<Variable>

Some other optional parameters accept variables for different purposes.

1. Baseline covariates of the longitudinal response: LongiCovars=<Variables list>
2. Variables with time-interaction for the longitudinal response: LongiTimeInter=<Variables list>
3. Baseline covariates of the time-to-event model: EventCovariates= <Variable list>
4. Stratification factors identifier for the time-to-event model: EventStrata= <Variable>
5. Variables for shared-parameter interactions: SharedInteraction= <Variables list>

Notes:

- All variables must be numeric except for the subject identifier and the stratification factor that can be either character or numeric.
- The %JM macro does not accept class factors as covariates. In case of categorical variables, the user needs to create 0-1 indicator variables and use these in the %JM macro.

7.2. Vertical input dataset

The %JM macro also allows the use of vertical input data structure using the option DataStructure=VERTICAL. This structure is especially useful to fit recurrent events, left and interval censoring and allows a new parameter (EventTimeStart=<variable>) to provide the starting time of each row. This structure can also be used to include time-dependent covariates.

The vertical input dataset structure requires a variable called JMDist that should equal "LONGI" for the longitudinal responses and "EVENT" for the time-to-event responses.

Table 6. Dataset in vertical structure (aidsvertical.sas7bdat)

Id	Obstime	JMDist	CD4	FromTime	ToTime	Death
1		EVENT		0	6.13	0
1		EVENT		6.13	11.07	0
1		EVENT		11.07	12.53	0
1		EVENT		12.53	13.93	0
1		EVENT		13.93	15.97	0
1		EVENT		15.97	16.97	0
1	0	LONGI	10.7			
1	6	LONGI	8.4			
1	12	LONGI	9.4			
2		EVENT		0	6.13	0
2		EVENT		6.13	11.07	0
2		EVENT		11.07	12.53	0
2		EVENT		12.53	13.93	0
2		EVENT		13.93	15.97	0
2		EVENT		15.97	17.8	0
2		EVENT		17.8	19	0
2	0	LONGI	6.3			
2	6	LONGI	8.1			
2	12	LONGI	4.6			
2	18	LONGI	5.1			

7.3. The parameters dataset

The macro saves the parameter estimates in a dataset chosen by the user in the OutputParameters parameter (or in work. jmoutputparameters by default) . The structure of this dataset is illustrated in Table 7.

Table 7. Parameters Dataset

Parameter	Estimate	StandardError	DF	tValue	Probt
L_INTERCEPT	7.2048	0.2218	465	32.49	<.0001
L_TIME_B1	-0.1921	0.02173	465	8.84	<.0001
L_TIME_B1_BY_DDI	0.01140	0.03026	465	0.38	0.7067
L_LOGSD_INTERCEPT	1.5223	0.03577	465	42.56	<.0001
L_COV_B0_B1	-0.03784	0.07188	465	0.53	0.5988
L_LOGSD_TIME_B1	-1.7003	0.09677	465	17.57	<.0001
L_LOGSD_RESIDUAL	0.5524	0.02749	465	20.09	<.0001
E_LOGHAZARD_1	-2.5160	0.1915	465	13.14	<.0001
E_LOGHAZARD_2	-2.2994	0.1788	465	12.86	<.0001
E_LOGHAZARD_3	-1.9222	0.2402	465	8.00	<.0001
E_LOGHAZARD_4	-2.5008	0.3423	465	7.31	<.0001
E_LOGHAZARD_5	-2.4188	0.3156	465	7.66	<.0001
E_LOGHAZARD_6	-2.4318	0.4018	465	6.05	<.0001
E_LOGHAZARD_7	-2.4121	0.5313	465	4.54	<.0001
E_DDI	0.3349	0.1565	465	2.14	0.0329
ASSOCT_TD	-0.2892	0.03603	465	-8.03	<.0001

A dataset with a "parameter" and an "estimate" column and parameters names as shown above is expected by the InitialParameters macro parameter in case users want force the use of a specific set of initial parameters. Sometimes it is useful to fit an initial model, save the parameters estimates in a

dataset using the OutputParameters macro parameter, and then use these estimates as initial parameters in subsequent calls to the macro using the InitialParameters macro parameter.

8. Validation Status

The %JM macro has been extensively tested but not followed a formal process of documented software validation [21]. The use of fully validated software may required in fields such as the analyses submitted to regulatory authorities (e.g. for the registration of a new drug).

This limitation may be addressed with one of the following strategies:

a) Use %JM macro, e.g. for sensitivity analysis, and add a warning note (e.g. in the statistical analysis plan): "The %JM macro is academic non-validated software. No other fully validated software (e.g. a supported SAS procedure) is available yet for this family of models."

b) Use the %JM macro as an one-off (single use) program and the JM R package as its corresponding QC program (or vice versa) to verify that both programs produce the same output. Since both the %JM macro and the JM R package were independently programmed (different programmers in a separate environment using distinct approximation and optimization routines), matching results provide strong evidence of a validated status for the output.

c) Program the model yourself for a specific model/data (single use program) and use the %JM macro as a QC program (or vice versa). You may contact A. Garcia-Hernandez & D. Rizopoulos for a tutorial on programming joint models for longitudinal and time-to-event data in SAS using PROC NLMIXED.

9. Macro History

The macro history is shown in Table 8.

Table 8. Macro History

Version	Date	Changes
01.01	15/12/2014	First Production Version
01.02	10/02/2015	Bug fixed in the detection of repeated knots. Bug fixed if piecewise was used with vertical. Two notes added, disjoint time-to-event model.
01.03	23/02/2015	Check for unexpected additional options
01.04	11/03/2015	Bug fixed in the creation of _JMpieces2 dataset Two new users notes added
01.05	29/04/2015	Bug fixed in the checking of SharedLongiTerm
01.06	01/05/2015	Name of _JMParameters_Normalized changed to _JMInitialParametersNormalized, for clarity . Oparameter renamed as NNestimate for clarity.
01.07	09/05/2015	Check that input dataset does not have variables with names used by the macro
01.08	24/08/2015	Several updates to avoid %SYSMACEXIST and make the macro workable before SAS 9.3. The NLMixedExpression macro parameter was added.
02.01	10/09/2015	- Two new possible associations have been added: BASELINE_VALUE and CHANGE_FROM_BASELINE. - The current value term can be included in the log-risk function after a transformation such as $f(x)=((x-10)**2)$ using the %transform_current_value_term macro and the &transform_current_value_term macro parameter.
02.02	28/10/2016	Bug in Poisson longitudinal response solved. A new example program (example7RM.sas) added to the zip file with a model fitting a Poisson distribution. Macro parameter called Random_Effects_in_Input_data added

10.   Bibliography

1. Tsiatis A. Davidian M. Joint modeling of longitudinal and Time-to-event data: an overview. *Statistica Sinica* 14(2004), 809-834.
2. Wu L, Liu W, Yi GC, Huang T. Analysis of longitudinal and survival Data: joint modeling , inference methods, and issues.
3. Rizopoulos D. Joint Models for Longitudinal and Time-to-Event Data: With Applications in R. Chapman & Hall/CRC Biostatistics Series. 2012.
4. Boyes ME. JMFit: A SAS Macro for Joint Models of Longitudinal and Survival. Datahttp://www.amstat.org/meetings/jsm/2015/onlineprogram/AbstractDetails.cfm?abstractid=314295
5. Rizopoulos D. JM R package. http://cran.r-project.org/web/packages/JM/
6. Rizopoulos D. JM: An R Package for the Joint Modeling of Longitudinal and Time-to-Event Data. *Journal of Statistical Software*. July 2010, Volume 35, Issue 9.
7. Philipson P, Diggle P, Sousa I, Kolamunnage-Dona R, Williamson P, Henderson R (2012). joineR: Joint modelling of repeated measurements and time-to-event data."
8. Proust-Lima C, Liquet B (2009). lcmm: an R package for estimation of latent class mixed models and joint latent class models." R cran.
9. Crowther MJ (2012). \STJM11: Stata module to fit shared parameter joint models of longitudinal and survival data." Statistical Software Components.
10. Rizopoulos D (2014). The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data using MCMC." arXiv preprint arXiv:1404.7625.
11. SAS/STAT® 9.2 User's Guide. The GLIMMIX Procedure. https://support.sas.com/documentation/cdl/en/statugglmmix/61788/PDF/default/statugglmmix.pdf
12. Farouki RT, Goodman TNT. On the optimal stability. Mathematics of Computation 65. 1996.
13. Farouki RT, Rajan VT. On the numerical condition of polynomials. Comput. Aided Geom. Design 4. 1987.
14. Racine J. A primer o regression splines. CRS R package. http://cran.r-project.org/web/packages/crs/
15. Hsieh F, Tseg YK, Wang JL. Joint modeling of survival and longitudinal data: likelihood approach revisited. 2006. Biometrics 62, 1037-1043.
16. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects." Statistics in medicine, 21(15), 2175-2197.
17. Abrams D, Goldman A, Launer C, Korvick J, Neaton J, Crane L, Grodesky M, Wakefield S, Muth K, Kornegay S, et al. (1994). \Comparative trial of didanosine and zalcitabine in patients with human immunodeficiency virus infection who are intolerant of or have failed zidovudine therapy." New England Journal of Medicine, 330, 657-662.
18. Murtaugh PA, Dickson ER, Van Dam GM, Malinchoc M, Grambsch PM, Langworthy AL, Gips CH. Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits. Hepatology. 1994 Jul;20(1 Pt 1):126-34.

19. Lesaffre E, Spiessens B (2001). On the effect of the number of quadrature points in a logistic random effects model: an example." Journal of the Royal Statistical Society: Series C (Applied Statistics), 50(3), 325-335.

20. Breslow NE, Lin X (1995). Bias correction in generalised linear mixed models with a single component of dispersion." Biometrika, 82(1), 81-91.

21. U.S. Department Of Health and Human Services. Food and Drug Administration. Center for Devices and Radiological Health. Center for Biologics Evaluation and Research *General Principles of Software Validation; Final Guidance for Industry and FDA Staff* http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126955.pdf