Table of Contents

1. Introduction ........................................................................................................................................ 4

2. Joint modeling of longitudinal and time-to-event data .......................................................... 4
   2.1. The longitudinal submodel ......................................................................................................... 5
       2.1.1. Modeling trajectories over time ............................................................................................ 6
       2.1.2. Random effects variance-covariance structures ................................................................. 8
       2.1.3. Parameterizations of the G matrix ....................................................................................... 9
   2.2. The time-to-event submodel ......................................................................................................... 9
       2.2.1. Specifying the baseline hazard ............................................................................................. 10
       2.2.2. Stratification ......................................................................................................................... 12
       2.2.3. Exogenous time-dependent covariates ............................................................................... 12
   2.3. Shared parameters ...................................................................................................................... 12
       2.3.1. A current-value dependent shared parameter ........................................................................ 12
       2.3.2. A slope dependent shared parameter .................................................................................... 13
       2.3.3. Cumulative effects ............................................................................................................... 13
       2.3.4. Lagged effects ...................................................................................................................... 13
       2.3.5. Random-effects coefficients ............................................................................................... 14
       2.3.6. Multiple shared-parameters ................................................................................................. 14
       2.3.7. Interaction effects ............................................................................................................... 14
       2.3.8. Association term .................................................................................................................. 15
   2.4. Obtaining the maximum-likelihood estimates ............................................................................ 15
       2.4.1. Conditional likelihood function ............................................................................................ 15
       2.4.2. Marginal likelihood function ............................................................................................... 17
       2.4.3. Optimization routines ......................................................................................................... 18
       2.4.4. Initial parameters ............................................................................................................... 18

3. Compiling and using the macro ....................................................................................................... 20
   3.1. Compiling the macro .................................................................................................................. 20
   3.2. Calling the macro ........................................................................................................................ 20

4. List of macro parameters .................................................................................................................. 21
   4.1. Identification of input data ......................................................................................................... 21
   4.2. Identification of output ............................................................................................................... 21
4.3. Set-up of the longitudinal model ................................................................. 22
4.4. Set-up of the time-to-event model ............................................................... 24
4.5. Set-up of the joint model ........................................................................... 25
4.6. Other parameters ....................................................................................... 27
5. JM package versus %JM macro: a comparative overview ........................................... 29
6. Examples ....................................................................................................... 30
   6.1. Example 1: a simple current-value shared-parameter model ......................... 30
   6.2. Example 2: a model with non-linear trajectories and interaction effects ........ 32
   6.3. Example 3: a model with a binary longitudinal response ............................ 34
   6.4. Example 4: improving the approximation of the survival function with vertical input data .... 37
   6.5. Example 5: a model with two shared-parameters: baseline-value and (trajectory) change-from-baseline .................................................................................. 39
   6.6. Example 6: a model with transformed (trajectory) current-value shared parameter .... 41
   6.7. Example 7: a model with a Poisson longitudinal response ......................... 43
   6.8. Example 8: Calculating dynamic predictions ........................................... 44
7. Input and output datasets .............................................................................. 46
   7.1. Standard input dataset .............................................................................. 46
   7.2. Vertical input dataset ............................................................................... 46
   7.3. The parameters dataset ........................................................................... 47
8. Validation Status .......................................................................................... 48
9. Macro History ............................................................................................... 49
10. Bibliography ............................................................................................... 50
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 Gompertz 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
\[
p(T_i, \delta_i, y_i | b_i; \theta) = p(T_i, \delta_i | b_i; \theta)p(y_i | b_i; \theta) \tag{2.1}
\]

where \(T_i\) is the observed time-to-event or time-to-censoring, \(\delta_i\) 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 \(\theta\) the vector of parameters that includes parameters for the time-to-event response \(\theta_T\) and for the longitudinal model \(\theta_y\). Under the assumptions of this model, the marginal (joint) log-likelihood of the observations on subject \(i\) is thus expressed as follows:

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

where \(\theta_b\) 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.

\[
p(y_i | b_i; \theta_Y) = \prod_j p(y_{ij}(t_{ij}) | b_i; \theta_y) \tag{2.3}
\]

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

\[
g\{m_i(t)\} = g[E\{y_i(t) | b_i\}] = X_i(t)(\beta_t + b_i) + Z_i b_b \tag{2.4}
\]

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 \(\theta_t\) and \(b_i\) respectively, and \(Z_i\) are \(\theta_b\) 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.

$$g\{m_i(t)\} = \beta_0 + \beta_1 t + \beta_2 \text{GENDER}_i + \beta_3 t \text{TRT}_i + b_{i0} + b_{i1} t, \quad (b_{i0}, b_{i1}) \sim N\left[\begin{array}{c} 
\sigma_0^2 \\
\sigma_0 \sigma_1 \\
\sigma_1^2 
\end{array}\right] \quad [2.5]$$

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>).

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

$$(\beta_{3+1} + b_{i3+1})(t - v_1)^3 + \cdots + (\beta_{3+N} + b_{i3+N})(t - v_N)^3 + Z_i\beta_b$$

where the expression $(\cdot)^+$ is used to indicate that negative values are set to zero, $v_1$ to $v_N$ are N internal knots, and $Z_i$ are $\theta_b$ are the design matrix and the coefficients associated with the baseline covariates. If
internal knots are not given by the user \(\text{Longiknot1}=<\text{time}>, \text{Longiknot2}=<\text{time}>, \text{etc}\), %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

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

where \(N\) is the number of internal knots, \(n\) the degree, and \(B_0(x)=1, B_1(x)=t, B_2(x)=t^2, \ldots, B_n(x)=t^n, B_{n+1}(x)=(t-\upsilon_1)^1, \ldots, B_{n+N}(x)=(t-\upsilon_N)^n\), 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 \((\text{LongiDegree}=2)\) B-spline with two internal knots \((\text{LongiNKnots}=2)\) 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 \((\text{LongiTimeInteraction}=\text{TRT})\) the %JM macro creates five basis functions variables \(B_0(x)\) to \(B_4(x)\). This model requires nine fixed-effect parameters, \(\beta_0\) to \(\beta_8\), and five random-effects parameters, \(b_0\) to \(b_4\), to fit the evolution of the longitudinal responses over time as shown below,

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

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 \((\text{LongiTimeModel} = \text{NATURALCUBIC})\) 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:

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

where \(N\) is the number of internal knots, \(d_j(t) = (t - \upsilon_j)^3/\left((\upsilon_{N+1} - \upsilon_j)^3\right)\) for \(j\) indexed from 1 to \(N\), \(\upsilon_0\) is the lower boundary knot, \(\upsilon_1\) to \(\upsilon_N\) are \(N\) internal knots and \(\upsilon_{N+1}\) refers to the upper boundary knot. If not given by the user, \(\upsilon_0\) is set to zero and \(\upsilon_{N+1}\) 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 \((\text{LongiGMMatrix} = \text{UN})\), \(k(k+1)/2\) 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.

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

With a variance-components option \((\text{LongiGMMatrix} = \text{VC})\), the \(G\) matrix is diagonal and only \(k\) variance parameters are required.

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

If an ante-dependence matrix is chosen \((\text{LongiGMMatrix} = \text{ANTE(1)})\), \(2k-1\) variance parameters are needed.

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

<table>
<thead>
<tr>
<th>Variance Component</th>
<th>LongiGMatrix=UN or VC</th>
<th>LongiGMatrix=ANTE(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LongiGMatrixParam=VARCOVAR</td>
<td>CHOL</td>
</tr>
<tr>
<td>( \sigma_0^2 )</td>
<td>( \text{EXP}(2\times\text{L_LOGSD_INTERCEPT}) )</td>
<td>( \text{L_GCHOL_B_1_1} \times \text{L_GCHOL_B_1_1} )</td>
</tr>
<tr>
<td>( \sigma_{01} )</td>
<td>( \text{L_COV_B0_B1} )</td>
<td>( \text{L_GCHOL_B_2_1} \times \text{L_GCHOL_B_1_1} )</td>
</tr>
<tr>
<td>( \sigma_1^2 )</td>
<td>( \text{EXP}(2\times\text{L_LOGSD_TIME_B1}) )</td>
<td>( \text{L_GCHOL_B_2_1} \times \text{L_GCHOL_B_1_1}+\text{L_GCHOL_B_2_2} \times \text{L_GCHOL_B_2_2} )</td>
</tr>
<tr>
<td>( \sigma_{02} )</td>
<td>( \text{L_COV_B0_B2} )</td>
<td>( \text{L_GCHOL_B_3_1} \times \text{L_GCHOL_B_1_1} )</td>
</tr>
<tr>
<td>( \sigma_{12} )</td>
<td>( \text{L_COV_B1_B2} )</td>
<td>( \text{L_GCHOL_B_3_1} \times \text{L_GCHOL_B_1_1}+\text{L_GCHOL_B_2_1} \times \text{L_GCHOL_B_2_1} )</td>
</tr>
<tr>
<td>( \sigma_2^2 )</td>
<td>( \text{EXP}(2\times\text{L_LOGSD_TIME_B2}) )</td>
<td>( \text{L_GCHOL_B_3_1} \times \text{L_GCHOL_B_1_1}+\text{L_GCHOL_B_2_1} \times \text{L_GCHOL_B_1_1}+\text{L_GCHOL_B_2_2} \times \text{L_GCHOL_B_2_2} )</td>
</tr>
</tbody>
</table>

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

\[
h_i(t) = h_0(t) \exp(y^T \omega_i) \tag{2.13}
\]

where \( \omega_i^T = (\omega_{i1}, ..., \omega_{i1}) \) denotes the vector of covariates (EventCovariates=<Variables list> ) and \( y \) 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):

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

**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).

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

**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:

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

where $\upsilon_0 = 0, \upsilon_1, ..., \upsilon_{Q-1}$ are Q internal knots that split the range of times into Q intervals, $\upsilon_{Q+1}$ is a value higher than the maximum observed time, and $\xi_q$ is the hazard at the interval defined from $\upsilon_{q-1}$ to $\upsilon_q$. 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 \( \kappa \) the vector of coefficients of the smooth function (EventModel= GSPLINE|GBSPLINE|GNATURALCUBIC). For example, a generalization using natural cubic splines (EventModel= GNATURALCUBIC) has the following formulation

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

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 \( \kappa \) the vector of coefficients of the smooth function (EventModel= WSPLINE|WBSPLINE|WNATURALCUBIC).

**Royston-Parmar model**

The Weibull model can also be 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 \( k_1 > 0 \). 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 \( \kappa \) 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

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

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

\[ h_i(t) = h_0(t) \exp\{y^T_1 \omega_i(t)\} \]  \[2.19\]

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:

\[ h_i(t) = h_0(t) \exp\{y^T \omega_i + \alpha m_i(t)\} \]  \[2.20\]
For the case that the baseline risk $h_0(t)$ is set as exponential, it should be noted that, once a shared-parameter $\alpha$ 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 $m_i(t) = m_i(0) + \text{ch}_m_i(t)$ where $\text{ch}_m_i(t) = m_i(t) - m_i(0)$.

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

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

where $\text{ch}_m_i(t) = m_i(t) - m_i(0)$ 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).

$$h_i(t) = h_0(t) \exp\{\gamma^T \omega_i + \alpha m'_i(t)\} \quad [2.21]$$

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).

$$h_i(t) = h_0(t) \exp[\gamma^T \omega_i + \alpha g(m_i(t))] \quad [2.22]$$

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>).
\[ h_i(t) = h_0(t) \exp\{y^T \omega_i + \alpha m_i[\max(t - c, 0)]\} \quad [2.23] \]

where \( c \) specifies the time lag of interest.

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

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

### 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 (\texttt{SharedParam= COEFFICIENTS}). If, for example, a random intercepts and slopes model is used for the longitudinal response the following model can be fitted:

\[ h_i(t) = h_0(t) \exp\{y^T \omega_i + \alpha_1 b_{i0} + \alpha_2 b_{i2}\} \quad [2.25] \]

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

### 2.3.6. Multiple shared-parameters

Several parameterizations, separated between blank spaces in the \texttt{SharedParam} parameter, can combined leading to multiple shared parameters \( \alpha \). For example, \texttt{SharedParam= CURRENT_VALUE SLOPE} would lead to the following model

\[ h_i(t) = h_0(t) \exp\{y^T \omega_i + \alpha_1 m_i(t) + \alpha_2 m_i'(t)\} \quad [2.26] \]

Including multiple parameterizations should be done cautiously. For example it is possible to include both a time-dependant shared-parameter \( \alpha_1 \) and a cumulative-dependant parameters \( \alpha_2 \), 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 (\texttt{SharedInteraction=<var list>}).
For example, SharedInteraction=TRT would change the basic time-dependant model of section 2.3.1 into the following model:

\[ h_i(t) = h_0(t) \exp\{\gamma^T \omega_i + \alpha_1 m_i(t) + \alpha_2 m_i(t) \cdot \text{TRT}\} \]  [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 \( m(t) \) as shown in equation 2.20 (\( \text{SharedLongiTerm} = \text{MEAN} \)); and b) the log-hazard function linearly related to the linear predictor \( g(m(t)) = X_i(t) + Z_i(t)\beta_i \) as illustrated in the equation below (\( \text{SharedLongiTerm} = \text{LINEAR\_PREDICTOR} \)).

\[ h_i(t) = h_0(t) \exp\{\gamma^T \omega_i + \alpha g(m_i(t))\} = h_0(t) \exp\{\gamma^T \omega_i + \alpha X_i(t)(\beta_i + b_i) + Z_i\beta_i\} \]  [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
\[
\log\{p(y_i(t_{ij}) | b_i; \theta_j)\} = -\frac{\log(2\pi)}{2} - \log(\sigma_e) - \frac{1}{2} \frac{[y_i(t_{ij})-m_{ij}]^2}{\sigma_e^2} \quad [2.29]
\]

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

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

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

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:

\[
\int_0^{T_i} h_i(t | b_i; \theta, \beta) dt \approx \sum_{k=1}^{15} w_k h_i(t_k | b_i; \theta, \beta) \quad [2.31]
\]

where \(t_k\) 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 (\texttt{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

<table>
<thead>
<tr>
<th>SubjectID</th>
<th>TimeFrom</th>
<th>TimeTo</th>
<th>EventYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>12</td>
<td>1</td>
</tr>
</tbody>
</table>

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

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

[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, $T_{i,0} = 0$ 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:

$$\log p(T_i, \delta_i, y_i | \theta_T, \theta_Y) = \log \int p(T_i, \delta_i | b_i; \theta_T, \theta_Y) \left[ \prod_i p(y_i(t_{ij}) | b_i; \theta_Y) \right] p(b_i; \theta_b) db_i$$

[2.33]

where $T_i$ is the observed time-to-event or time-to-censoring, $\delta_i$ is the indicator of event or censoring, $y_i(t_{ij})$ the collection of observations in the longitudinal response for subject $i$, $\theta_T$ and $\theta_Y$ the vectors of parameters for the time-to-event and the longitudinal model respectively, and $\theta_b$ 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>
<thead>
<tr>
<th>Optimization Technique</th>
<th>NLMIXEDOptions</th>
<th>Time (minutes)</th>
<th>-2LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Quasi-Newton</td>
<td>tech=quanew</td>
<td>11:36</td>
<td>8656.8</td>
</tr>
<tr>
<td>(or left blank)</td>
<td>tech=dbldog</td>
<td>8:04</td>
<td>8656.8</td>
</tr>
<tr>
<td>Newton-Raphson with search</td>
<td>tech=newrap</td>
<td>17:28</td>
<td>8656.8</td>
</tr>
<tr>
<td>Newton-Raphson with ridging</td>
<td>tech=nrridg</td>
<td>17:42</td>
<td>8656.8</td>
</tr>
<tr>
<td>Trust region</td>
<td>tech=trureg</td>
<td>17:32</td>
<td>8656.8</td>
</tr>
</tbody>
</table>

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 $\alpha$, 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);
```

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 $y(t)$. 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 fit 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>

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 | WBSPPLINE | 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>
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 \( m(t) \) [or \( m'(t) \)]. 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 \( g(m(t)) \).

- **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:

- **CALCULATEEXECETIME**: 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.

```markdown
%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 b0, b1, b2, 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

<table>
<thead>
<tr>
<th></th>
<th>JM R package V1.2.0</th>
<th>%JM SAS macro V 1.04</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Longitudinal Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trajectories over time: linear, polynomials, splines, B-splines, natural cubic splines</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Baseline Covariates</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Time-interaction terms</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Time-to-Event Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PH Weibull</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PH Piecewise-exponential</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PH Unsspecified baseline</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>PH Generalization of Gompertz based on splines</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PH Generalization of Weibull based on splines</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>PH Royston-Parmar model</td>
<td>X (^5)</td>
<td>X</td>
</tr>
<tr>
<td>AFT Weibull</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Exogenous time-dependent covariates</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Stratification Factors</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Parameterizations to link both models</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time-dependant (trajectory Model)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Slope-dependant</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Cumulative-dependant</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Lagging</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Random-effects coefficients</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Interaction effects</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Approximation of the survival function</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7-points Gauss-Kronrod Rule</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>15-points Gauss-Kronrod Rule</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Approximation of the integral over the random effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gauss-Hermitage</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Adaptive Gauss-Hermitage</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pseudo-adaptive Gauss-Hermitage</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Laplace Approximation</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Optimization Techniques</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EM algorithm</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Quasi-Newton algorithm</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Other methods available in NL MIXED</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

\(^5\) 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 $b_{i1}$.

$$m_i(t) = \beta_0 + \beta_1 t + \beta_2 t \text{DDI}_i + b_{i0} + b_{i1} t$$  \hfill [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

$$h_i(t) = h_0(t) \exp\{\gamma_1 \text{DDI}_i + \alpha m_i(t)\}$$  \hfill [6.1.2]

For such a model the SAS code is given by

```sas
%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  = CALCULATEEXECUTE TIME);  
```
For this model, the %JM macro provides the following output:

Joint Model Summary:
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 -3030.82
... ... ... ... ...
41 73 4328.38921 4.64E-11 1.279E-6 -889E-13

NOTE: FCNV 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

| Label                                      | Estimate | Standard 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 2        | -2.2994  | 0.1788         | 465 | -12.86  | <.0001|   |
| Event Process: Log Hazard Period 3        | -1.8222  | 0.2402         | 465 | -7.31   | <.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        | -2.4211  | 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:

\[
m_i(t) = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t + (\beta_2 + b_{i2})t^2 + \beta_3 D_{PNCi} + \beta_4 t D_{PNCi} + \beta_5 t^2 D_{PNCi} \quad [6.2.1]
\]
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:

\[
h_i(t) = h_0(t) \exp[\gamma_1 D_{PNC} + \gamma_2 HMG + \alpha_1 m_i(t) + \alpha_2 (m_i(t) \times HMG)]\] [6.2.2]

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

```latex
%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 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 b10 b11 b12
Distribution for Random Effects Normal
Subject Variable id
Optimization Technique Dual Quasi-Newton
Integration Method Adaptive Gaussian
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 $\pi_i$, 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 \( b_{i1} \).

\[
\log \left( \frac{\pi_i(t)}{1-\pi_i(t)} \right) = \beta_0 + \beta_1 t + \beta_2 t \text{DDI}_i + b_{i0} + b_{i1} t \quad [6.3.1]
\]

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

\[
h_i(t) = h_0(t) \exp\{\gamma_1 \text{DDI}_i + \alpha \pi_i(t)\} \quad [6.3.2]
\]

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,
    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
| Label                                      | Estimate | Standard 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: DDD 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  | 1.2416         | 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|
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

```sas
%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 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: b10 b11
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

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>4373.34496</td>
<td>15.73627</td>
<td>102.4166</td>
<td>-3038.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>70</td>
<td>4328.59044</td>
<td>3.21E-10</td>
<td>0.000012</td>
<td>-629E-12</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Label</th>
<th>Estimate</th>
<th>Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitudinal Process: Intercept</td>
<td>7.2045</td>
<td>0.2218</td>
<td>465</td>
<td>32.49</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: Slope</td>
<td>-0.1920</td>
<td>0.0217</td>
<td>465</td>
<td>-8.84</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: DDI by slope</td>
<td>0.01150</td>
<td>0.03027</td>
<td>465</td>
<td>0.38</td>
<td>0.7041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Residual</td>
<td>0.5524</td>
<td>0.02749</td>
<td>465</td>
<td>20.09</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Random Intercept</td>
<td>1.5224</td>
<td>0.03577</td>
<td>465</td>
<td>42.56</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Random Slope</td>
<td>-1.7001</td>
<td>0.09673</td>
<td>465</td>
<td>-17.57</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal Process: Covariance Intercept by Slope</td>
<td>-0.03823</td>
<td>0.07189</td>
<td>465</td>
<td>-0.53</td>
<td>0.5951</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 1</td>
<td>-2.5221</td>
<td>0.1915</td>
<td>465</td>
<td>-13.17</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 2</td>
<td>-2.2872</td>
<td>0.1785</td>
<td>465</td>
<td>-12.81</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 3</td>
<td>-1.9595</td>
<td>0.2403</td>
<td>465</td>
<td>-8.15</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 4</td>
<td>-2.4993</td>
<td>0.3413</td>
<td>465</td>
<td>-7.32</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 5</td>
<td>-2.4142</td>
<td>0.3157</td>
<td>465</td>
<td>-7.65</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 6</td>
<td>-2.4005</td>
<td>0.4009</td>
<td>465</td>
<td>-5.99</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 7</td>
<td>-2.4221</td>
<td>0.5302</td>
<td>465</td>
<td>-4.57</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Process: DDI</td>
<td>0.3333</td>
<td>0.1565</td>
<td>465</td>
<td>2.13</td>
<td>0.0337</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association (current-value)</td>
<td>-0.2888</td>
<td>0.03602</td>
<td>465</td>
<td>-8.02</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 $b_{i1}$.

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

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

$$h_i(t) = h_0(t) \exp\{\gamma_1 DDI_i + \alpha_1 m_i(0) + \alpha_2 \cdot m_i(t)\} \quad [6.5.2]$$

For such a model the SAS code is given by

```sas
%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 = CALCULATEXECTIME);
```
For this model, the %JM macro provides the following output:

Joint Model Summary:
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

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

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

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 shown in example 1. An unstructured covariance was assumed for the G matrix so an extra parameter captures the covariance between $b_{i0}$ and $b_{i1}$.

$$m_1(t) = \beta_0 + \beta_1 t + \beta_2 t DDI_i + b_{i0} + b_{i1}t$$ \[6.6.1\]

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:

$$h_i(t) = h_0(t) \exp \left\{ \gamma_1 DDI_i + \alpha \{m_1(t) - 10\}^2 \right\}$$ \[6.6.2\]

For such a model the SAS code is given by

```sas
%macro transform_current_value_term(cvt=);
((&cvt.- 10)**2)
%mend;

%mJM( 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=CALCULATEEXECUTE TIME, 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**

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>4376.4675</td>
<td>12.54504</td>
<td>1007.156</td>
<td>-101294</td>
</tr>
<tr>
<td>59</td>
<td>131</td>
<td>4332.54132</td>
<td>2.06E-10</td>
<td>0.000891</td>
<td>-0.02241</td>
</tr>
</tbody>
</table>

**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**

<table>
<thead>
<tr>
<th>Label</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitudinal Process: Intercept</td>
<td>7.2049</td>
<td>0.2217</td>
<td>465</td>
<td>32.50</td>
</tr>
<tr>
<td>Longitudinal Process: DD by slope</td>
<td>-0.2076</td>
<td>0.02182</td>
<td>465</td>
<td>-9.51</td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Residual</td>
<td>0.006229</td>
<td>0.03046</td>
<td>465</td>
<td>0.02241</td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Random Intercept</td>
<td>1.5225</td>
<td>0.0373</td>
<td>465</td>
<td>42.61</td>
</tr>
<tr>
<td>Longitudinal Process: Log SD Random Slope</td>
<td>-1.6613</td>
<td>0.09209</td>
<td>465</td>
<td>-18.04</td>
</tr>
<tr>
<td>Longitudinal Process: Covariance Intercept by Slope</td>
<td>0.01118</td>
<td>0.07242</td>
<td>465</td>
<td>0.15</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 1</td>
<td>-5.3032</td>
<td>0.2574</td>
<td>465</td>
<td>-20.60</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 2</td>
<td>-5.1594</td>
<td>0.2773</td>
<td>465</td>
<td>-18.61</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 3</td>
<td>-4.8074</td>
<td>0.3256</td>
<td>465</td>
<td>-14.76</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 4</td>
<td>-5.3913</td>
<td>0.4078</td>
<td>465</td>
<td>-13.22</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 5</td>
<td>-5.3416</td>
<td>0.3992</td>
<td>465</td>
<td>-13.38</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 6</td>
<td>-5.3775</td>
<td>0.4810</td>
<td>465</td>
<td>-11.18</td>
</tr>
<tr>
<td>Event Process: Log Hazard Period 7</td>
<td>-5.3059</td>
<td>0.6149</td>
<td>465</td>
<td>-8.63</td>
</tr>
<tr>
<td>Event Process: DDI</td>
<td>0.3282</td>
<td>0.1774</td>
<td>465</td>
<td>1.85</td>
</tr>
<tr>
<td>Association (current-value)</td>
<td>0.03204</td>
<td>0.003429</td>
<td>465</td>
<td>9.35</td>
</tr>
</tbody>
</table>

**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 \( \pi_i \).

*In this dataset we create cd4 as Poisson assuming it is more or less a Poisson variable with offset called offset_var;*

```sas
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 \( b_{i1} \).

\[
\log\{m_i(t)\} = \beta_0 + \beta_1 t + \beta_2 t \text{DDI}_i + b_{i0} + b_{i1} t \quad [6.7.1]
\]

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

\[
h_i(t) = h_0(t) \exp\{\gamma_1 \text{DDI}_i + \alpha m_i(t)\} \quad [6.7.2]
\]

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

```sas
%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,
    LongiVar = 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.

<table>
<thead>
<tr>
<th>Patient</th>
<th>ddi</th>
<th>Time</th>
<th>Death</th>
<th>obstime</th>
<th>cd4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>1</td>
<td>6.1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 2</td>
<td>0</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject 3</td>
<td>0</td>
<td>12.5</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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:

$$\pi(u|t) = \Pr\{T_i^* \geq u | T_i^* > t; Y_i(t); \theta\} \approx \frac{S_i(u|\hat{M}_i(u\hat{\theta}_i(t));\hat{\theta})}{S_i(t|\hat{M}_i(u\hat{\theta}_i(t));\hat{\theta})}$$

[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>
<thead>
<tr>
<th>Id</th>
<th>Obstime</th>
<th>CD4</th>
<th>Time</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>10.7</td>
<td>16.97</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>8.4</td>
<td>16.97</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>9.4</td>
<td>16.97</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>6.3</td>
<td>19.00</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>8.1</td>
<td>19.00</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>4.6</td>
<td>19.00</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>5.0</td>
<td>19.00</td>
<td>0</td>
</tr>
</tbody>
</table>

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)

<table>
<thead>
<tr>
<th>Id</th>
<th>Obstime</th>
<th>jMDist</th>
<th>CD4</th>
<th>FromTime</th>
<th>ToTime</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EVENT</td>
<td>0</td>
<td>6.13</td>
<td>0</td>
<td>6.13</td>
<td>11.07</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>11.07</td>
<td>12.53</td>
<td>0</td>
<td>12.53</td>
<td>13.93</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>13.93</td>
<td>15.97</td>
<td>0</td>
<td>15.97</td>
<td>16.97</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>15.97</td>
<td>16.97</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>0</td>
<td>6.13</td>
<td>0</td>
<td>6.13</td>
<td>11.07</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>6.13</td>
<td>11.07</td>
<td>0</td>
<td>12.53</td>
<td>13.93</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>11.07</td>
<td>12.53</td>
<td>0</td>
<td>15.97</td>
<td>17.8</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>15.97</td>
<td>17.8</td>
<td>0</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>EVENT</td>
<td>17.8</td>
<td>19</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>LONGI</td>
<td>10.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>LONGI</td>
<td>8.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12</td>
<td>LONGI</td>
<td>9.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

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

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>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>01.01</td>
<td>15/12/2014</td>
<td>First Production Version</td>
</tr>
<tr>
<td>01.02</td>
<td>10/02/2015</td>
<td>Bug fixed in the detection of repeated knots.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bug fixed if piecewise was used with vertical.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two notes added, disjoint time-to-event model.</td>
</tr>
<tr>
<td>01.03</td>
<td>23/02/2015</td>
<td>Check for unexpected additional options</td>
</tr>
<tr>
<td>01.04</td>
<td>11/03/2015</td>
<td>Bug fixed in the creation of _JMpieces2 dataset Two new users notes added</td>
</tr>
<tr>
<td>01.05</td>
<td>29/04/2015</td>
<td>Bug fixed in the checking of SharedLongTerm</td>
</tr>
<tr>
<td>01.06</td>
<td>01/05/2015</td>
<td>Name of _JMParameeters_Normalized changed to _JMIniitalParamatersNormalized, for clarity. Oparameter renamed as NNestimate for clarity.</td>
</tr>
<tr>
<td>01.07</td>
<td>09/05/2015</td>
<td>Check that input dataset does not have variables with names used by the macro</td>
</tr>
<tr>
<td>01.08</td>
<td>24/08/2015</td>
<td>Several updates to avoid %SYSMACEXIST and make the macro workable before SAS 9.3. The NLMixedExpression macro parameter was added.</td>
</tr>
<tr>
<td>02.01</td>
<td>10/09/2015</td>
<td>- Two new possible associations have been added: BASELINE_VALUE and CHANGE_FROM_BASELINE.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- 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 &amp;transform_current_value_term macro parameter.</td>
</tr>
<tr>
<td>02.02</td>
<td>28/10/2016</td>
<td>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.</td>
</tr>
</tbody>
</table>
10. Bibliography

