

A flexible joint-modeling framework for longitudinal and time-to-event data with overdispersion

Peer-reviewed author version

NJAGI, Edmund; MOLENBERGHS, Geert; Rizopoulus, Dimitris; VERBEKE, Geert; Kenward, Michael G.; DENDALE, Paul & WILLEKENS, Koen (2013) A flexible joint-modeling framework for longitudinal and time-to-event data with overdispersion. In: STATISTICAL METHODS IN MEDICAL RESEARCH, Published online before print July 18, 2013, p. 1-16.

DOI: 10.1177/0962280213495994

Handle: <http://hdl.handle.net/1942/15819>

# A Flexible Joint Modelling Framework for Longitudinal and Time-to-event Data with Overdispersion

Edmund Njeru Njagi<sup>1\*</sup> Geert Molenberghs<sup>1,2</sup> Dimitris Rizopoulos<sup>3</sup>  
Geert Verbeke<sup>2,1</sup> Michael G. Kenward<sup>4</sup>  
Paul Dendale<sup>5,6</sup> Koen Willekens<sup>7</sup>

<sup>1</sup> *I-BioStat, Universiteit Hasselt, B-3590 Diepenbeek, Belgium*

<sup>2</sup> *I-BioStat, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium*

<sup>3</sup> *Department of Biostatistics, Erasmus University Medical Center, NL-3000 CA Rotterdam, the Netherlands*

<sup>4</sup> *Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London WC1E7HT, United Kingdom*

<sup>5</sup> *Hasselt University, Faculty of Medicine, B-3590 Diepenbeek, Belgium*

<sup>6</sup> *Jessa Hospital, Heart Centre Hasselt, B-3500 Hasselt, Belgium*

<sup>7</sup> *Katholieke Universiteit Leuven, Faculty of Medicine, B-3000 Leuven, Belgium*

## Abstract

We combine conjugate and normal random effects in a joint model for outcomes, at least one of which is non-Gaussian, with particular emphasis on cases in which one of the outcomes is of survival type. Conjugate random effects are used to relax the often-restrictive mean-variance prescription in the non-Gaussian outcome, while normal random effects account for not only the correlation induced by repeated measurements from the same subject, but the association between the different outcomes. Using a case study in chronic heart failure, we show that model fit can be improved, even resulting in impact on significance tests, by switching to our extended framework. By first taking advantage of the ease of analytical integration over conjugate random effects, we easily estimate our framework, by maximum likelihood, in standard software.

*Some Keywords:* Partial marginalization, Poisson-gamma-normal model, Probit-beta-normal model, Weibull-gamma-normal model.

## 1 Introduction

The collection of outcomes of different types from study subjects is common in research. In HIV/AIDS studies, apart from the time to onset of AIDS or death, viral load and CD4 cell count may be recorded repeatedly over time. In cancer studies, several biological markers may be collected over time, together with the time to death or metastasis. An example is in prostate cancer, where, after treatment

---

\*Address for correspondence: I-BioStat, Universiteit Hasselt, Agoralaan 1, B-3590 Diepenbeek, Belgium. Email: edmund.njagi@uhasselt.be

for cancer, prostate-specific antigen measurements may be collected over time, and the time to disease recurrence recorded<sup>1,2</sup>. Andersen et al.<sup>3</sup> describe a study in liver cirrhosis in which apart from the survival time, several biochemical variables were collected at follow-up, among them bilirubin, albumin, and prothrombin. Moreover, the time-to-event outcome may also be a repeated outcome. This is the case when it is recurrent, or when it is collected from subjects within clusters. Dendale et al.<sup>4</sup> and Njagi et al.<sup>5</sup> describe a study in cardiology, in which researchers, through telemonitoring (a process through which patients are remotely monitored), not only repeatedly measured daily blood pressure, heart rate and weight from initially discharged chronic heart failure patients, but also recorded the time-to-rehospitalization. Time-to-rehospitalization in this case was a recurrent survival outcome, since a discharged patient could be rehospitalized more than once over time. Duchateau and Janssen<sup>6</sup> describe a study, in veterinary research, where the protein and ureum concentrations were repeatedly measured over time in dairy cows, and the time to first insemination was also recorded. Since cows were clustered within herds, the time-to-event was a repeated outcome.

The objectives thereof may be varied, and some may require a joint modelling approach. Interest may be on the distribution of one of the outcomes, conditional on the other. For instance, the distribution of the time-to-event, conditional on the history of the “true” longitudinal biomarker process<sup>7–9</sup>. It may also be of interest whether covariates have an effect on all the outcomes jointly. Furthermore, research questions do not necessarily have to be restricted to the outcomes in their original form; rather, they may be directed at modified versions of the outcomes, where, say, objectives require a joint analysis of a dichotomized version of an originally continuous longitudinal outcome, and, say, a survival outcome. In this respect, there therefore is no limit as to the possible combinations of different types of outcomes that the statistician is bound to encounter.

Now, on the one hand, though not in the context of joint modelling, Molenberghs et al.,<sup>10</sup> in building on the work by Molenberghs, Verbeke and Demétrio<sup>11</sup> paid emphasis on models for binomial, count, and survival data. It is known that such models exhibit a mean-variance prescription. Random effects with a conjugate distribution are usually incorporated in order to relax this prescription, whenever data at hand violate it. Note that such violation is referred to as overdispersion. This leads, for instance, to the beta-binomial and the negative-binomial models for binomial and count data respectively. Whenever these data are obtained within a hierarchical setting, common practice is to use models incorporating normally distributed random effects, in the form of generalized linear mixed models, to account for correlation among repeated measurements from the same experimental

unit. The above-mentioned authors highlight that since both overdispersion and hierarchies may occur simultaneously, it would be befitting to consider a model in which, rather than make a choice for either the conjugate or the normal random effects, the strength of both is combined in a single model. In this way, the “combined” model explicitly pays attention to the three structures that exist in the context of hierarchical non-Gaussian outcomes: the mean structure, the variance structure, and the correlation structure. Therefore, simultaneously, the often-restrictive mean-variance relationship is relaxed through conjugate random effects, and the correlation induced by hierarchies is accounted for through normal random effects. Through case studies, the authors show that model fit can be substantially improved by switching to the “combined” framework. This may even have impact on hypothesis tests for covariate effects. Even when interest is on simpler (less flexible) models, the extended framework can still be used as a goodness-of-fit tool.

Turning attention to joint modelling, on the other hand, the shared-parameter modelling framework<sup>7;8</sup> is a popular approach in the so-called field of joint modelling of longitudinal and time-to-event data. It entails postulating a submodel for the longitudinal outcome, a submodel for the time-to-event, and linking the two submodels through a shared latent structure. The latent structure can but does not have to be parametric. In the parametric case, for instance, shared normal random effects are usually considered; these play the role of taking into account the correlation between the longitudinal measurements, and the association between the time-to-event and longitudinal outcomes<sup>8</sup>. The field has, and is currently, experiencing rapid growth, and excellent reviews are provided by Tsiatis and Davidian<sup>7</sup> and Yu et al.<sup>1</sup> It is evident that most of the work in this field has focussed on cases in which the time-to-event is univariate. It is also evident that most of the focus has been on applications in which the time-to-event is analyzed jointly with a continuous longitudinal outcome. Finally, it is evident that whenever a parametric choice for the shared latent structure is made, in which case it is mostly of the Normal type, focus is only on accounting for the correlation induced by repeated measures in the longitudinal outcome, and the association between the longitudinal and the time-to-event outcome. It is instructive to reiterate here that even univariate survival outcomes can exhibit overdispersion<sup>10</sup>.

In this paper, we present an extended shared random effects framework for the joint analysis of pairs of outcomes of different types: time-to-event and continuous; time-to-event and binary; time-to-event and count. We also consider an additional case of continuous and binary outcomes. We pay attention to the possibility that each of the different outcomes may be obtained as repeated measures. The

framework can easily be extended to analyze any number of different types of outcomes. The flexibility of our framework stems from the fact that in formulating the shared-parameter model, each submodel for a non-Gaussian outcome incorporates both conjugate and normal random effects. Therefore, each submodel for a non-Gaussian outcome reflects all the structures exhibited by repeated measures non-Gaussian outcomes, as highlighted earlier. This work therefore builds on Molenberghs et al.<sup>10</sup> to make shared-parameter joint modelling more flexible. Through a case study, we show that the additional flexibility can provide improvement to model fit, and even have impact on significance tests. Our choice for conjugate random effects is based on the attendant ease of analytical integration, which we exploit to easily estimate the model via partial marginalization<sup>11</sup>. Nevertheless, we show that it's possible to derive closed form-expressions for the fully marginalized joint probabilities; indeed, estimation could proceed by specifying the marginal likelihood contributions, though existence of infinite series may render the same cumbersome.

The organization of the paper is as follows: In the following section, we introduce the chronic heart failure data. This will be followed by a brief review of the *Weibull-gamma-normal*, *probit-beta-normal*, and the *Poisson-gamma-normal* models<sup>10</sup>, which constitute the ingredients for our framework. In Section 4, we introduce our modeling framework, followed by a discussion on estimation in Section 5. The data are analyzed in Section 6, and the analysis program is provided in the Appendix.

## 2 Case Study: The Chronic Heart Failure Data

These data originated from a study conducted in Belgium between 2008 and 2010, and whose objective was to study whether follow-up of chronic heart failure (CHF) patients, by means of a telemonitoring program, reduced mortality and rehospitalization rates<sup>4</sup>. Heart failure is a condition in which the heart has difficulty pumping enough blood for the needs of the body. Daily measurements of systolic and diastolic blood pressure, heart rate and weight, were remotely collected from 80 patients. These patients, at hospital discharge, were provided with a set of apparatuses, through which they not only made these measurements, but also remotely availed the measurements to the medical personnel. These longitudinal measurements were recorded each day for a period of about 6 months. 16 patients had rehospitalizations; 13 of these once, 2 twice, and 1 thrice. In addition, the following patient characteristics were also collected at baseline: patient's sex, age, heart rhythm, NTproBNP as a measure of cardiac muscle fiber stretch, patient fitness indicator (given using the New York Heart Association, NYHA, class indication), and the Left Ventricular Ejection Fraction

(LVEF), which is a measure of heart performance. We will jointly model the risk of rehospitalization and the mean number of times a patient's heart rate measurements could be classified as "abnormal", with LVEF as a baseline covariate. Notice that the survival outcome is repeated (recurrent).

### 3 Review of Ingredients

In this section, we briefly review the extended generalized linear modelling framework for repeated measures, discussed by Molenberghs et al.<sup>10</sup>, and which combines conjugate and normal random effects. We only highlight the specific cases of the *Weibull-gamma-normal*, *probit-beta-normal*, and the *Poisson-gamma-normal* models. We also mention the linear mixed model, for completeness purposes, and since it will, together with the aforementioned, be an ingredient for our framework.

#### 3.1 Weibull-gamma-normal Model

For repeated measures time-to-event outcomes, the above-mentioned present the following, for the  $k^{th}$  survival time in cluster  $i$ ,  $k = 1, \dots, p_i$ :

$$f(t_{ik}|\psi_{ik}, \mathbf{b}_i) = \lambda \rho t_{ik}^{\rho-1} \psi_{ik} e^{\mu_{ik} + d_{ik}} e^{-\lambda_k t_{ik}^{\rho} \psi_{ik} e^{\mu_{ik} + d_{ik}}}, \quad (1)$$

$$\psi_{ik} \sim \text{Gamma}(\alpha_k, \beta_k), \quad (2)$$

$$d_{ik} = \mathbf{w}'_{ik} \mathbf{b}_i, \mathbf{b}_i \sim N(0, D), \quad (3)$$

where  $\mu_{ik} = \mathbf{x}'_{ij} \boldsymbol{\xi}$ , where  $\boldsymbol{\xi}$  is a vector of fixed effects parameters. Forthwith, whenever  $\mathbf{b}_i$  is used, the above distributional assumption will apply. The baseline hazard function is therefore assumed to follow the parametric form leading to Weibull distribution for the event times. The gamma random effects, of conjugate distribution to the Weibull, are assumed independent. This is not too restrictive, as correlation between the repeated measures is induced by normal random effects. However, dependence between the gamma random effects can also be assumed. One could also consider a common gamma random effect for a cluster:

$$\psi_i \sim \text{Gamma}(\alpha, \beta).$$

### 3.2 Probit-beta-normal Model

For repeated measures binary outcomes, the authors consider both the logistic-beta-normal and the probit-beta-normal. The distinction is in usage of the logit link for the former, and the probit for the latter. Here, however, we only review the latter. Closed form expressions are possible with the latter, and in case interest is on the former, there exists back-transformation approximations. With  $Y_{ij}$  the  $j^{th}$  measurement in cluster  $i$ ,  $j = 1, \dots, n_i$ , they consider the following:

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij}), \quad (4)$$

$$\pi_{ij} \sim \theta_{ij} \Phi(\mathbf{x}'_{ij} \boldsymbol{\gamma} + \mathbf{z}'_{ij} \mathbf{b}_i), \quad (5)$$

$$\theta_{ij} \sim \text{Beta}(\alpha_1, \beta_1), \quad (6)$$

where  $\Phi(\cdot)$  is the cumulative distribution function of the normal distribution.  $\boldsymbol{\gamma}$  is a vector of fixed effects parameters. The model combines both normal ( $\mathbf{b}_i$ ) and conjugate (beta) random effects. The vectors  $\mathbf{x}'_{ij}$  and  $\mathbf{z}'_{ij}$  are rows stacked into the design matrices  $X_i$  and  $Z_i$ , corresponding to the fixed and random effects respectively. This will hold whenever these vectors are used.

### 3.3 Poisson-gamma-normal Model

For repeated measures count data, the authors consider, with  $Y_{ij}$ ,  $j = 1, \dots, n_i$  the  $j^{th}$  count in the  $i^{th}$  cluster, the following model:

$$\begin{aligned} P(Y_{ij} = y_{ij} | \theta_{ij}, \mathbf{b}_i) &= \frac{1}{y_{ij}!} \left( \theta_{ij} e^{\tau_{ij} + \mathbf{z}'_{ij} \mathbf{b}_i} \right)^{y_{ij}} e^{-\theta_{ij} e^{\tau_{ij} + \mathbf{z}'_{ij} \mathbf{b}_i}}, \\ \theta_{ij} &\sim \text{Gamma}(\alpha_j, \beta_j), \\ \tau_{ij} &= \mathbf{x}'_{ij} \boldsymbol{\zeta}, \end{aligned} \quad (7)$$

where  $\boldsymbol{\zeta}$  is a vector of fixed effects parameters. Both normal ( $\mathbf{b}_i$ ) and conjugate (gamma) random effects are incorporated; as in the Weibull-gamma-normal case, various settings for the gamma random effect could be considered.

### 3.4 Linear Mixed Model

Models for continuous outcomes, just like those for survival, binary, and count data, are not only brought together in the generalized linear modelling framework, but also in the exponential family

mathematical framework. The mean and variance are separate parameters in the case of continuous normally distributed outcomes. Random effects with a normal distribution, which provide a conjugate distribution to the normal linear model, are incorporated to induce intracluster correlation in repeated measures continuous outcomes. For the  $j^{\text{th}}$  measurement in cluster  $i$ ,  $j = 1, \dots, n_i$ ,

$$\mathbf{Y}_{ij} | \mathbf{b}_i \sim N(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i, \sigma^2) \quad (8)$$

is the hierarchically specified linear mixed model, where  $\boldsymbol{\beta}$  is a vector of fixed effects parameters.

#### 4 Flexible Joint Modelling Framework

In the foregoing section, we have highlighted the *Weibull-gamma-normal*, *probit-beta-normal*, and the *Poisson-gamma-normal* framework, discussed by Molenberghs et al.<sup>10</sup>. The framework explicitly addresses each of the structures characteristic of hierarchical non-Gaussian outcomes: the mean, the variance, and the correlation structure. The mean-variance prescription is relaxed, and correlation is accounted for, using two different sets of random effects. One of the advantages of switching to this extended framework is the gain in terms of model fit. This can also have impact on tests for covariate effects. The ease of analytical integration over conjugate random effects also makes the framework easy to estimate by maximum likelihood, via partial marginalization<sup>11</sup>. Partial marginalization involves analytical integration over the conjugate random effects, leaving the normal random effects for numerical integration, using such a tool as the SAS procedure NL MIXED. Finally, even when interest is on simpler (less flexible) models, the extended framework can still be used as a goodness-of-fit tool.

We utilize this framework to develop an extended and flexible shared random effects joint modelling framework. By allowing the submodels for non-Gaussian outcomes to reflect all the structures exhibited by repeated measures non-Gaussian outcomes, through the incorporation of both conjugate and normal random effects, our work aims at providing a framework which provides improvement to model fit. In the following, we will consider specific cases of pairs of different types of repeated measures outcomes, though extension to any number of different types of outcomes is straightforward, as mentioned earlier.



#### 4.1 Case 1: Repeated Time-to-event and Repeated Continuous Outcomes

For the continuous outcome submodel, we formulate the linear mixed model, as in (8), for the  $j^{th}$  measurement in cluster  $i$ ,  $j = 1, \dots, n_i$ . We reiterate here that  $\mathbf{b}_i$  follows the distribution as stated in Section 3.1, which will hold throughout.

For the time-to-event outcome submodel, we formulate the Weibull-gamma-normal model, specified by (1), (2), and (3), for the  $k^{th}$  survival time in cluster  $i$ ,  $k = 1, \dots, p_i$ . The  $k$  gamma random effects are assumed independent, with parameters  $\alpha$  and  $\beta$ . This could be relaxed, for example, to allow for serial association; however, this possibility will not be pursued here. Moreover,  $d_{ik} = \mathbf{w}'_{ik} \mathbf{b}_i$ , where  $\mathbf{w}'_{ik}$  is a vector of scale factors. For the scale and shape parameters in the baseline hazard, we consider a more general case, where both  $\lambda$  and  $\rho$  are allowed to vary between members of a cluster.

The continuous and survival processes are assumed independent, conditional on the shared normal random effects. The joint model, conditional on both the normal and gamma random effects, thus takes the form

$$f(\mathbf{t}_i, \mathbf{y}_i | \mathbf{b}_i, \boldsymbol{\psi}_i) = \prod_k \lambda_k \rho_k t_{ik}^{\rho_k - 1} \psi_{ik} e^{\mu_{ik} + d_{ik}} e^{-\lambda_k t_{ik}^{\rho_k} \psi_{ik} e^{\mu_{ik} + d_{ik}}} \times \frac{1}{(2\pi)^{\frac{n_i}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2} (\mathbf{y}_i - X_i \boldsymbol{\beta} - Z_i \mathbf{b}_i)' \Sigma_i^{-1} (\mathbf{y}_i - X_i \boldsymbol{\beta} - Z_i \mathbf{b}_i)}, \quad (9)$$

with  $\Sigma_i$  an  $n_i$  by  $n_i$  diagonal covariance matrix with diagonal elements  $\sigma^2$ . Note that the shared random effect in the way considered here is generic. For example, one can choose  $\mathbf{z}'_{ij}$  and  $\mathbf{w}'_{ij}$  such that some random effects are present in the normal-outcome linear predictors, with others influencing the Weibull predictor, and a third set influencing both. As such, our paradigm encompasses both shared as well as correlated random effects.

#### 4.2 Case 2: Repeated Time-to-event and Repeated Binary Outcomes

For the binary outcome submodel, we formulate the Probit-beta-normal model, given in (4), (5), and (6), for the  $j^{th}$  measurement in cluster  $i$ ,  $j = 1, \dots, n_i$ . For the time-to-event outcome submodel, we formulate the Weibull-gamma-normal model, with all the specifications as in Section 4.1. Conditional independence of the binary and survival processes, given the shared normal random effects, is again assumed. By first noting that integrating out the beta random effects from the model specified by

(4), (5), and (6) can be shown to imply that

$$f(y_{ij}|\mathbf{b}_i) = \frac{1}{\alpha_1 + \beta_1} (K_{ij}\alpha_1)^{y_{ij}} [(1 - K_{ij})\alpha_1 + \beta_1]^{1-y_{ij}}, \quad (10)$$

where

$$K_{ij} = \Phi(\mathbf{x}'_{ij}\boldsymbol{\gamma} + \mathbf{z}'_{ij}\mathbf{b}_i), \quad (11)$$

the form of the joint conditional model can then be seen as

$$\begin{aligned} f(\mathbf{y}_i, \mathbf{t}_i|\mathbf{b}_i, \boldsymbol{\psi}_i) &= \prod_j \frac{1}{\alpha_1 + \beta_1} (K_{ij}\alpha_1)^{y_{ij}} [(1 - K_{ij})\alpha_1 + \beta_1]^{1-y_{ij}} \\ &\times \prod_k \lambda_k \rho_k t_{ik}^{\rho_k - 1} \psi_{ik} e^{\mu_{ik} + d_{ik}} e^{-\lambda_k t_{ik}^{\rho_k} \psi_{ik} e^{\mu_{ik} + d_{ik}}}. \end{aligned} \quad (12)$$

The conditioning here is only on the gamma and normal random effects, given that the beta random effects in the binary outcome submodel has been integrated out.

### 4.3 Case 3: Repeated Time-to-event and Repeated Count Outcomes

For the repeated counts submodel, we formulate the Poisson-gamma-normal model, given in (7), while for the time-to-event outcome submodel, we consider the Weibull-gamma-normal model, with specifications once again as in Section 4.1. Conditional independence of the two processes is again assumed, given the shared normal random effects. The joint model, conditional on the random effects, then takes the form

$$\begin{aligned} P(\mathbf{Y}_i = \mathbf{y}_i, \mathbf{T}_i = \mathbf{t}_i|\mathbf{b}_i, \boldsymbol{\Theta}_i, \boldsymbol{\Psi}_i) &= \prod_j \frac{1}{y_{ij}!} \left( \theta_{ij} e^{\tau_{ij} + \mathbf{z}'_{ij}\mathbf{b}_i} \right)^{y_{ij}} e^{-\theta_{ij} e^{\tau_{ij} + \mathbf{z}'_{ij}\mathbf{b}_i}} \\ &\times \prod_k \lambda_k \rho_k t_{ik}^{\rho_k - 1} \psi_{ik} e^{\mu_{ik} + d_{ik}} e^{-\lambda_k t_{ik}^{\rho_k} \psi_{ik} e^{\mu_{ik} + d_{ik}}}. \end{aligned} \quad (13)$$

Note that here we have two sets of gamma random effects; one set in the survival, and the other in the count process.

### 4.4 Case 4: Repeated Binary and Repeated Continuous Outcomes

Finally, we consider this commonly encountered setting. For the continuous outcome submodel, we consider the linear mixed model, as specified in Section 4.1. For the binary outcome submodel, we

consider the Probit-beta-normal model specified in (4), (5), and (6). Conditional independence, as described in the foregoing sections, is once again invoked. For notational distinction between the binary and the continuous outcome,  $\mathbf{Y}_{1i}$  is used for the continuous outcome, and  $\mathbf{Y}_{2i}$  is used for the binary outcome. Precisely, for the binary outcome, letting  $Y_{2ik}$  be the  $k^{th}$  measurement in cluster  $i$ ,  $k = 1, \dots, p_i$ , we assume:

$$\begin{aligned} Y_{2ik} &\sim \text{Bernoulli}(\pi_{ik}), \\ \pi_{ik} &\sim \theta_{ik} \Phi(\mathbf{x}'_{ik} \boldsymbol{\gamma} + \mathbf{z}'_{ik} \mathbf{b}_i), \\ \theta_{ik} &\sim \text{Beta}(\alpha_1, \beta_1). \end{aligned}$$

Again, notice that by integrating out the beta random effects, these model specifications for the binary process can be shown to imply that

$$f(y_{2ik} | \mathbf{b}_i) = \frac{1}{\alpha_1 + \beta_1} (M_{ik} \alpha_1)^{y_{2ik}} [(1 - M_{ik}) \alpha_1 + \beta_1]^{1 - y_{2ik}}, \quad (14)$$

where

$$M_{ik} = \Phi(\mathbf{x}'_{ik} \boldsymbol{\gamma} + \mathbf{z}'_{ik} \mathbf{b}_i).$$

The joint model for the continuous and the binary process, this time only conditional on the normal random effects, is then

$$\begin{aligned} f(\mathbf{y}_{2i}, \mathbf{y}_{1i} | \mathbf{b}_i) &= \prod_k \frac{1}{\alpha_1 + \beta_1} (M_{ik} \alpha_1)^{y_{2ik}} [(1 - M_{ik}) \alpha_1 + \beta_1]^{1 - y_{2ik}} \\ &\times \frac{1}{(2\pi)^{\frac{n_i}{2}} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} e^{-\frac{1}{2} [\mathbf{y}_{1i} - (X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i)]' \boldsymbol{\Sigma}_i^{-1} [\mathbf{y}_{1i} - (X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i)]}. \end{aligned} \quad (15)$$

## 5 Estimation and Inference

It is possible to derive the fully marginalized joint model probabilities, and therefore parameter estimation for our framework can proceed through maximum likelihood, by directly specifying the marginal likelihood contributions. As an illustration, we derive this for both Case 1 and Case 3 above. The relevant calculations are briefly highlighted in the Appendix. While the marginal probabilities could be directly specified for estimation, the existence of infinite series, as can be seen from the expressions, may make the approach intractable. In all apart from the binary-continuous case, they contain

infinite series. However, given the ease of analytical integration over the conjugate random effects, and the availability of software that can numerically integrate over normal random effects, analytical integration and numerical integration can be combined to provide a convenient estimation route. This is the idea behind the partial marginalization approach of Molenberghs, Verbeke and Demétrio<sup>11</sup>. It involves first analytically integrating out the conjugate random effects, and leaving the normal random effects for numerical integration, such as implemented in the SAS procedure NLMIXED. Therefore, for estimation, we require only the expressions for the joint distribution marginal over the conjugate but conditional on the normal random effects. In the following, we provide expressions for the partially marginalized models, for our four cases presented above.

**Case 1: Repeated Time-to-event and Repeated Continuous Outcomes:**

We consider the model (9), and integrate over the gamma random effects. We then have the following as the joint distribution conditional on the normal random effects:

$$f(\mathbf{t}_i, \mathbf{y}_i | \mathbf{b}_i) = \frac{1}{(2\pi)^{\frac{n_i}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2} [\mathbf{y}_i - (X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i)]' \Sigma_i^{-1} [\mathbf{y}_i - (X_i \boldsymbol{\beta} + Z_i \mathbf{b}_i)]} \times \prod_k \frac{\lambda_k \rho_k t_{ik}^{\rho_k - 1} e^{\mu_{ik} + d_{ik}} \times \alpha}{\left( \lambda_k t_{ik}^{\rho_k} e^{\mu_{ik} + d_{ik}} + \frac{1}{\beta} \right)^{\alpha+1} \times \beta^\alpha}. \quad (16)$$

**Case 2: Repeated Time-to-event and Repeated Binary Outcomes:**

Considering model (12), and integrating over gamma random effects, leads to the joint distribution conditional on the normal random effects:

$$f(\mathbf{t}_i, \mathbf{y}_i | \mathbf{b}_i) = \prod_j \frac{1}{\alpha_1 + \beta_1} (K_{ij} \alpha_1)^{y_{ij}} [(1 - K_{ij}) \alpha_1 + \beta_1]^{1 - y_{ij}} \times \prod_k \frac{\lambda_k \rho_k t_{ik}^{\rho_k - 1} e^{\mu_{ik} + d_{ik}} \times \alpha}{\left( \lambda_k t_{ik}^{\rho_k} e^{\mu_{ik} + d_{ik}} + \frac{1}{\beta} \right)^{\alpha+1} \times \beta^\alpha},$$

with  $K_{ij}$  as in (11).

### Case 3: Repeated Time-to-event and Repeated Count Outcomes:

We integrate over both  $\Theta_i$  and  $\Phi_i$  from model (13), having

$$f(\mathbf{t}_i, \mathbf{y}_i | \mathbf{b}_i) = \prod_j \left( \frac{1}{y_{ij}!} \right) \left( e^{\tau_{ij} + \mathbf{z}'_{ij} \mathbf{b}_i} \right)^{y_{ij}} \frac{\Gamma(y_{ij} + \alpha_j)}{\Gamma(\alpha_j) \beta_j^{\alpha_j}} \left( e^{\tau_{ij} + \mathbf{z}'_{ij} \mathbf{b}_i} + \frac{1}{\beta_j} \right)^{y_{ij} + \alpha_j} \\ \times \prod_k \lambda_k \rho_k t_{ik}^{\rho_k - 1} e^{\mu_{ik} + d_{ik}} \frac{\alpha_k}{\beta_k^{\alpha_k} \left( \lambda_k t_{ik}^{\rho_k} e^{\mu_{ik} + d_{ik}} + \frac{1}{\beta_k} \right)^{\alpha_k + 1}} \quad (17)$$

as the joint model, conditional on normal random effects.

### Case 4: Repeated Binary and Repeated Continuous Outcomes:

For this case, (15) already provides the required expression, given that the beta random effect has already been marginalized out of the model for the binary outcome.

These expressions are all that is required to use, for example, the SAS procedure NLMIXED for estimation and inference.

Note that though we here focus on maximum likelihood estimation, other estimation methods can also be considered; for instance, a Bayesian approach.

## 6 Analysis of the Chronic Heart Failure Data

In this analysis, we jointly model the recurrent time-to-rehospitalization and a count version of the dichotomized longitudinal heart rate. Understanding rehospitalization is important in heart failure management. Heart rate was first dichotomized into “normal” (50–90; coded 0) and “abnormal” (values higher than 90; coded 1). Values less than 50 are not considered for this analysis. During each period in which a patient was not under hospitalization, the number of times that the patient’s heart rate measurements were classified as “abnormal” was enumerated, generating a count response. Notice that patients who were rehospitalized and discharged at least once in the course of the study had at least 2 periods in which they were not under hospitalization, separated by a period of hospitalization. As such, the count outcome was also a repeated response. As a covariate, we considered the baseline left ventricular ejection fraction, LVEF. LVEF indicates the fraction of blood being pumped out of the ventricle with each contraction. We consider two categories for this covariate; ( $>45\%$ ; coded 0), and ( $\leq 45\%$ ; coded 1), forthwith referred to as “preserved ejection,” and “reduced ejection,” respectively.

For the time-to-rehospitalization submodel, we considered

$$f_{ik}(t|\psi_i, b_i) = \lambda \rho t^{\rho-1} \psi_i e^{\mu_i + d_i} e^{-\lambda(t_{2ik}^\rho - t_{1ik}^\rho) \psi_i e^{\mu_i + d_i}}, \quad (18)$$

for the  $k^{th}$  risk period of patient  $i$ , where  $d_i = \kappa b_i$ ,  $\kappa$  a scale factor, and  $b_i \sim N(0, \sigma_b^2)$ . Note that  $t_{1ik}$  and  $t_{2ik}$  respectively represent the beginning and the end of the risk period, where  $t_{1ik} \leq t \leq t_{2ik}$ . Notice that representation of the at-risk periods was done using the calendar-time format<sup>6</sup>. Further,  $\mu_i = X_i \xi$ ; and  $X_i = 1$  if subject had reduced ejection, and 0 otherwise. The gamma random effect,  $\psi_i$ , was assumed to follow a one-parameter distribution:

$$\psi_i \sim \text{Gamma}(\alpha, \alpha^{-1}), \quad (19)$$

for identifiability purposes<sup>6</sup>.

For the repeated heart rate counts submodel, on the other hand, with  $Y_{ij}$  being the number of times subject  $i$  exhibited abnormal heart rate measurements during the  $j^{th}$  period in which the subject was not under hospitalization, we considered

$$P(\mathbf{Y}_i = \mathbf{y}_i | \theta_i, b_i) = \prod_j \frac{1}{y_{ij}!} \left( \theta_i e^{\tau_{ij} + b_i} \right)^{y_{ij}} e^{-\theta_i e^{\tau_{ij} + b_i}},$$

with  $\tau_{ij} = \log(\text{time}) + \zeta_0 + \zeta_1 X_i$ ,  $X_i$  representing the baseline left ventricular ejection status as before, and  $b_i \sim N(0, \sigma_b^2)$ . Notice the inclusion of an offset term,  $\log(\text{time})$ , to account for differing time periods within which the counts were made. Identifiability of both parameters of the gamma distribution in this context is problematic, as discussed by Molenberghs, Verbeke and Demétrio<sup>11</sup>. To circumvent this,  $\theta_i$  was assumed to follow a one-parameter distribution:

$$\theta_i \sim \text{Gamma}(\alpha_2, \alpha_2^{-1}). \quad (20)$$

In line with (17), and taking our specifications into account, the joint model for the survival and

count outcome, conditional upon the normal random effect, takes the form:

$$f(y_{ij}, t|b_i) = \prod_j \frac{1}{y_{ij}!} \left( e^{\tau_{ij}+b_i} \right)^{y_{ij}} \frac{\Gamma(y_{ij} + \alpha_2) \times \alpha_2^{\alpha_2}}{\Gamma(\alpha_2) \times (e^{\tau_{ij}+b_i} + \alpha_2)^{y_{ij}+\alpha_2}} \times \frac{\lambda \rho t^{\rho-1} e^{\mu_i+d_i} \times \alpha^{\alpha+1}}{\{\lambda(t_{ik2}^\rho - t_{ik1}^\rho) e^{\mu_i+d_i} + \alpha\}^{\alpha+1}}. \quad (21)$$

By assuming that the likelihood contribution of censored observations is that of survival probabilities, we allow censoring of the survival outcome in the following manner:

$$f(y_{ij}, t|b_i) = \prod_j \frac{1}{y_{ij}!} \left( e^{\tau_{ij}+b_i} \right)^{y_{ij}} \frac{\Gamma(y_{ij} + \alpha_2) \times \alpha_2^{\alpha_2}}{\Gamma(\alpha_2) \times (e^{\tau_{ij}+b_i} + \alpha_2)^{y_{ij}+\alpha_2}} \times \frac{(\lambda \rho t^{\rho-1} e^{\mu_i+d_i})^{\delta_i} \times \Gamma(\alpha + \delta_i) \times \alpha^\alpha}{\Gamma(\alpha) \left( \lambda(t_{ij2}^\rho - t_{ij1}^\rho) e^{\mu_i+d_i} + \alpha \right)^{\alpha+\delta_i}}, \quad (22)$$

where  $\delta_i$  is a censoring indicator, taking the value 1 if the observation is an event time, and 0 otherwise.

Implementation was done in the SAS procedure NLMIXED, and the analysis program is presented in the Appendix. Given that we need to define different likelihood contributions for the two different outcomes, we use the general likelihood feature, as provided for in the above procedure. This feature has also been utilized by Molenberghs and Verbeke,<sup>12</sup> in estimating joint models for continuous and discrete outcomes. Starting values were obtained by fitting a separate independent poisson and survival model to the count and survival outcome respectively. The results are summarized in Table 1. Alongside results from our extended joint model, results from what we refer to as the conventional model, are shown. The conventional model only incorporates the normal random effect, to account for both the correlation induced by repeated measurements from the same subject, as well as the association between the two responses, as opposed to our extended model which goes a step further to relax the mean-variance prescriptions in the respective submodels, through the inclusion of the gamma random effects, as described above. A 5% level of significance was used.

We first look at the results from our extended model. The test for a joint effect of ejection status on both processes is not statistically significant ( $p = 0.1650$ ), and therefore we conclude that there is no statistically significant evidence of a joint effect of ejection status on both the mean number of abnormal heart rate measurements and the risk of rehospitalization. Based on exponentiation of

the relevant parameter estimate, the mean number of abnormal heart rate measurements in patients with reduced ejection is found to be 3.3531 times that of patients with preserved ejection. This effect is at borderline statistical significance ( $p = 0.0594$ ). The risk of rehospitalization for patients with reduced ejection is obtained, by also exponentiating the corresponding parameter estimate, as 5.5168 times that of patients with preserved ejection; however, this effect is clearly not statistically significant ( $p = 0.6498$ ). Finally, the estimate of the scale factor,  $\kappa$ , is positive, implying that patients with a higher mean number of abnormal heart rate measurements have increased risk of rehospitalization. However, this should not be over-emphasized, given that it is not statistically significant ( $p = 0.3201$ ).

We now compare results from the extended and the conventional model. Based on an AIC-based comparison, we observe that our extended model provides improvement to model fit, without compromising parsimony. There is a noticeable impact on both the point estimates and standard errors. As noted above, the effect of ejection status on the mean number of abnormal heart rate measurements is borderline significant under the extended model; however, the case is quite different under the conventional model ( $p = 0.0901$ ). There is also a remarkable difference in the scale factor; it is highly significant under the conventional model ( $p = 0.0022$ ), while this is clearly not the case under the extended model, as mentioned above. However, in terms of the hypothesis of a joint effect of ejection status on both processes, the two models provide close results; ( $p = 0.1650, 0.1648$ ) for the extended and the conventional model respectively. It is important to recall that in univariate generalized linear models for non-Gaussian outcomes, overly restrictive variance functions pose the risk of incorrect standard errors<sup>13</sup>. Though the joint modelling case is different since more outcomes are involved, it would be expected that too parsimonious variance structures may pose similar problems. This will be investigated in follow-up simulation studies, to study the impact of omitting the conjugate random effects on specific model parameters, under different scenarios.

Finally, we consider the predictive ability of the extended model as compared to that of the conventional model, with respect to the time-to-event outcome. In assessing predictive performance, we use the approach of concordance, as described by Harrell et al.<sup>14</sup>. This approach is based on first considering all pairs of patients, where at least one member of the pair has experienced the event of interest. If the event times in such a pair can be ordered, the pair is considered “usable.” The logic is then that for each “usable” pair, the predicted probability of surviving up to a certain point should be larger for the member of the pair who survives longer. Pairs for which this condition holds are



considered “concordant”, and the proportion of “usable” pairs for which concordance holds, provides the concordance index<sup>14</sup>.

The extended and the conventional models were both fitted to the first 90 days’ data. Based on the model estimates, and assuming a time zero starting point, the predicted probabilities of remaining rehospitalization-free for 90 days were computed, for each patient. These predictions were compared to the actual first rehospitalization events which occurred between day 91 and day 180, in order to provide the concordance index. Note that in constructing the “usable” pairs, we only consider the patient’s first rehospitalization event in the concerned period (the day 91-day 180 period).

Under the extended model, and once again assuming (19) for the distribution of  $\psi_i$ , the survival probabilities described above are of the form

$$\begin{aligned} P(T > t|b_i) &= \int_0^\infty [\exp\{-\int_0^t \lambda \rho s^{\rho-1} \psi_i \exp(\mu_i + d_i) ds\}] f(\psi_i) d\psi_i, \\ &= \frac{\alpha^\alpha}{\{\alpha + \lambda t^\rho \exp(\mu_i + d_i)\}}, \end{aligned} \quad (23)$$

where  $t = 90$ .

On the other hand, under the conventional model, they are of the form

$$P(T > t|b_i) = \exp\{-\lambda t^\rho \exp(\mu_i + d_i)\}. \quad (24)$$

Estimates for the probabilities were computed based on the maximum likelihood estimates of the respective model parameters. For the normal random effects, the empirical Bayes estimates were used.

A total of 434 pairs were usable. Based on the extended model, 328 of these were concordant, 105 discordant, and 1 pair had tied predicted survival probabilities. This therefore provided an index of

$$\frac{328 + 0.5(1)}{434} = 75.6912\%.$$

For the conventional model, on the other hand, 310 were concordant, 123 discordant, and 1 pair had tied probability predictions; this gave an index of

$$\frac{310 + 0.5(1)}{434} = 71.5438\%.$$

**Table 1:** *Chronic Heart Failure Data. Parameter estimates (standard errors) for the extended joint repeated counts/recurrent time-to-event model and for the conventional analysis.*

Effect	Parameter	Extended model	Conventional model
		Estimate (s.e.)	Estimate (s.e.)
Count process			
Intercept	$\zeta_0$	-3.9913 (0.8269)	-5.4085 (0.6501)
LVEF status	$\zeta_1$	1.2099 (0.6326)	1.2247 (0.7138)
Variance (Gamma random effects)	$1/\alpha_2$	2.7528 (1.2335)	
Normal random effects			
Variance	$\sigma_b^2$	1.3662 (1.6182)	5.4479 (1.2892)
Survival process			
LVEF status	$\xi$	1.7078 (3.7476)	0.0119 (0.5871)
Scale	$\lambda$	0.000045 (0.0004)	0.0047 (0.0056)
Shape	$\rho$	1.4718 (1.5600)	0.6998 (0.1937)
Variance (Gamma random effects)	$1/\alpha$	7.0263 (12.9652)	
Scale factor	$\kappa$	1.8304 (1.8294)	0.4212 (0.1333)
Joint effect of LVEF status (p-value)			
		0.1650	0.1648
Model fit [Akaike information criterion (AIC)]			
		792.7461	806.4592

Therefore, the extended model seems capable of better discriminating between patients who are going to be rehospitalized within the 90 days, and the ones who will not.

A useful extension that would enable a more comprehensive treatment to the subject of prediction would be to connect our work to that of dynamic prediction,<sup>9;15;16</sup> where conditional patient-specific probabilities of surviving later time points are computed from a fitted joint model, given the patient's available information up to a certain time point. The so-called dynamic discrimination indices are then computed. Actually, in the estimation of the probability of being rehospitalization free provided by (23), we partly utilize the previously recorded information of each patient by plugging in the empirical Bayes estimates of the random effects  $b_i$ , which are the modes of the posterior distribution of the random effects given the observed data of the patient. However, even in such a context, the calculation of discrimination indices in the context of recurrent events is not straightforward, and more work would be required to adequately address this.

## **7 Concluding Remarks**

We have outlined a very broad and extended, hence flexible, joint modelling framework. Apart from the correlation induced by repeated measurements from the same subject and the association between the different outcomes, the often-restrictive mean-variance prescription in the model for the non-Gaussian outcome has explicitly been addressed. This has been done conveniently through the inclusion of conjugate random effects, an aspect that has been exploited to easily estimate the framework in standard software, through partial marginalization.

Through our analyses of data from the area of chronic heart failure, we have shown that our extended framework provides improvement to model fit, while still maintaining parsimony, and also offers better prediction. We have also observed impact on significance tests.

Follow-up work will involve general assessments, through simulation studies, on the effects of omitting the conjugate random effects. More specifically, the impact on specific model parameters, under various scenarios, will be investigated. These scenarios include the level of overdispersion in each of the non-Gaussian outcomes, the amount of censoring, as well as the length of the longitudinal sequence.

This framework also opens avenues for further research work in related areas. It is possible to derive the marginal joint correlation functions, which may be of interest, for example, in surrogate marker evaluation and in psychometrics.

## **Acknowledgment**

We gratefully acknowledge support from IAP research Network P6/03 of the Belgian Government (Belgian Science Policy).

## A Appendix

### A.1 Derivations for the Joint Marginal Probabilities

#### A.1.1 Case 1: Repeated Time-to-event and Repeated Continuous Outcomes

Consider the conditional joint distribution given by (9). The following expansion

$$e^{-\lambda_k t_{ik} \rho_k \psi_{ik} e^{\mu_{ik} + d_{ik}}} = \sum_{\ell_k=0}^{\infty} \frac{(-1)^{\ell_k}}{\ell_k!} \lambda_k^{\ell_k} \psi_{ik}^{\ell_k} t_{ik}^{\ell_k \rho_k} e^{\ell_k \mu_{ik}} e^{\ell_k d_{ik}} \quad (25)$$

facilitates the integration process and, integrating over the normal random effects, the joint distribution, conditional on the gamma random effects, takes the form:

$$\begin{aligned} f(\mathbf{t}_i, \mathbf{y}_i | \boldsymbol{\psi}_i) &= \frac{1}{(2\pi)^{\frac{n_i}{2}} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} e^{-\frac{1}{2} (\mathbf{y}_i - X_i \boldsymbol{\beta})' \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - X_i \boldsymbol{\beta})} \\ &\times \sum_{\boldsymbol{\ell}} \frac{(-1)^{\sum_k \ell_k}}{\prod_k \ell_k!} e^{\sum_k (\ell_k + 1) \mu_{ik}} \left( \prod_k \psi_{ik}^{\ell_k + 1} \lambda_k^{\ell_k + 1} \rho_k t_{ik}^{\rho_k \ell_k + \rho_k - 1} \right) \times a_{\ell k} \\ &\times \frac{1}{|D|^{\frac{1}{2}}} \times |(D^{-1} + Z_i' \boldsymbol{\Sigma}_i Z_i)^{-1}|^{\frac{1}{2}}, \end{aligned} \quad (26)$$

where

$$a_{\ell k} = e^{\frac{1}{2} [\sum_k (\ell_k + 1) \mathbf{w}_{ik}' - 2(\mathbf{y}_i - X_i \boldsymbol{\beta})' \boldsymbol{\Sigma}_i^{-1} Z_i] [D^{-1} + Z_i' \boldsymbol{\Sigma}_i Z_i]^{-1} [\sum_k (\ell_k + 1) \mathbf{w}_{ik}' - 2(\mathbf{y}_i - X_i \boldsymbol{\beta})' \boldsymbol{\Sigma}_i^{-1} Z_i]'}$$

Integrating-out the gamma random effects, we get

$$\begin{aligned} f(\mathbf{t}_i, \mathbf{y}_i) &= \frac{1}{(2\pi)^{\frac{n_i}{2}} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} e^{-\frac{1}{2} (\mathbf{y}_i - X_i \boldsymbol{\beta})' \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - X_i \boldsymbol{\beta})} \\ &\times \sum_{\boldsymbol{\ell}} \frac{(-1)^{\sum_k \ell_k}}{\prod_k \ell_k!} e^{\sum_k (\ell_k + 1) \mu_{ik}} \left( \prod_k \lambda_k^{\ell_k + 1} \rho_k t_{ik}^{\rho_k \ell_k + \rho_k - 1} \right) \times a_{\ell k} \times b_{\ell k} \\ &\times \frac{1}{|D|^{\frac{1}{2}}} \times |(D^{-1} + Z_i' \boldsymbol{\Sigma}_i Z_i)^{-1}|^{\frac{1}{2}}, \end{aligned} \quad (27)$$

where  $b_{\ell k} = \{[\beta^{\ell_k - 1} (\ell_k + \alpha) (\ell_k + \alpha - 1) \dots (\alpha)]^{p_i}\}$ , and  $a_{\ell k}$  is as defined above. Further, summation over  $\boldsymbol{\ell}$  is shorthand for multi-index summation over all components of  $(\ell_1, \ell_2, \dots, \ell_{p_i})$ .

### A.1.2 Case 3: Repeated Time-to-event and Repeated Count Outcomes

Consider the conditional joint distribution given in (13). We make use of the expansion

$$e^{-\theta_{ij}} e^{\tau_{ij} + \mathbf{z}'_{ij} \mathbf{b}_i} = \sum_{m_j=0}^{\infty} \frac{(-1)^{m_j}}{m_j!} \theta_{ij}^{m_j} e^{m_j \tau_{ij}} e^{m_j \mathbf{z}'_{ij} \mathbf{b}_i}$$

and of expansion (25), and after integrating-out the normal random effects, the joint distribution, conditional on the two sets of gamma random effects, is

$$\begin{aligned} P(\mathbf{Y}_i = \mathbf{y}_i, \mathbf{T}_i = \mathbf{t}_i | \Theta_i, \Psi_i) &= \left( \prod_{j=1}^{n_i} \frac{1}{y_{ij}!} \right) \sum_{\mathbf{m}} \sum_{\mathbf{l}} (-1)^{\sum_j m_j + \sum_k \ell_k} \left( \prod_j \theta_{ij}^{m_j + y_{ij}} \right) \left( \prod_k \Psi_{ik}^{\ell_k + 1} \right) \\ &\times \prod_k \left( \lambda_k^{\ell_k + 1} \rho_k t_{ik}^{\rho_k \ell_k + \rho - 1} \right) \times e^{[\sum_j (m_j + y_{ij}) \tau_{ij} + \sum_k (\ell_k + 1) \mu_{ik}]} \\ &\times g_{m_j \ell_k}, \end{aligned} \quad (28)$$

where

$$g_{m_j \ell_k} = e^{\frac{1}{2} [\sum_j (m_j + y_{ij}) \mathbf{z}'_{ij} + \sum_k (\ell_k + 1) \mathbf{w}'_{ik}] D [\sum_j (m_j + y_{ij}) \mathbf{z}_{ij} + \sum_k (\ell_k + 1) \mathbf{w}_{ik}]}$$

Upon integrating-out the gamma random effects, we obtain the marginal joint distribution as

$$\begin{aligned} P(\mathbf{Y}_i = \mathbf{y}_i, \mathbf{T}_i = \mathbf{t}_i) &= \prod_{j=1}^{n_i} \frac{1}{y_{ij}!} \sum_{\mathbf{m}} \sum_{\mathbf{l}} (-1)^{\sum_j m_j + \sum_k \ell_k} \\ &\times \left[ \prod_j \frac{\beta_j^{m_j + y_{ij}} \Gamma(\alpha_j + m_j + y_{ij})}{\Gamma(\alpha_j)} \right] \left[ \prod_k \frac{\beta_k^{\ell_k + 1} \Gamma(\alpha_k + \ell_k + 1)}{\Gamma(\alpha_k)} \right] \\ &\times \prod_k \left( \lambda_k^{\ell_k + 1} \rho_k t_{ik}^{\rho_k \ell_k + \rho - 1} \right) \times e^{[\sum_j (m_j + y_{ij}) \tau_{ij} + \sum_k (\ell_k + 1) \mu_{ik}]} \\ &\times g_{m_j \ell_k} \end{aligned} \quad (29)$$

, with  $g_{m_j \ell_k}$  as defined above.

## A.2 Analysis programs

Our analysis program is provided below. Notice that the variable “outcome” brings together a patient’s measurements on both outcomes, with the variable “type” distinguishing between survival and count outcomes, in order to correspondingly specify the respective likelihood contributions.

```

proc nlmixed data=chf;
bounds rho>0, alpha>0, lambda>0, alpha2>0;
parms zeta0=-3.5165 zeta1=1.0288 xi=0.076599190 lambda=0.003375627 rho=0.809716599;
offset=1;
if type='outc1' then do;
eta=offset*logtime+zeta0+zeta1*(lvef=1)+b;
loglik=count*eta-lgamma(count+1)+lgamma(count+alpha2)-lgamma(alpha2)+alpha2*log(alpha2)
-(count+alpha2)*log(exp(eta)+alpha2);
end;
else if type='outc2' then do;
loglik=censor*log(lambda*rho*(day**(rho-1))*exp(xi*(lvef=1)+scalef*b))
+lgamma(alpha+censor)-log(((lambda*(rend**rho-rstart**rho)*
exp(xi*(lvef=1)+scalef*b)+alpha)**(alpha+censor))*((1/alpha)**alpha))-lgamma(alpha);
end;
model outcome ~ general(loglik);
random b ~ normal(0, sigmab**2) subject=ptid;
estimate 'Variance of Normal R.E.s' sigmab**2;
estimate 'Variance of Gamma R.E.s Survival' 1/alpha;
estimate 'Variance of Gamma R.E.s Count' 1/alpha2;
contrast 'Joint effect hypothesis' xi,zeta1;
run;

```

## References

1. Yu M, Law, N, et al. Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica* 2004; **14**: 835–832.
2. Menggang Y, Taylor JMG and Sandler HM. Individual Prediction in Prostate Cancer Studies Using a Joint Longitudinal SurvivalCure Model. *Journal of the American Statistical Association* 2008; **103**: 178–187.
3. Andersen P, Borgan O, et al. *Statistical Models Based on Counting Processes*. New York: Springer, 1993.
4. Dendale P, De Keulenaer G, et al. Effect of a telemonitoring-facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (TElemonitoring in the MAnagement of Heart Failure) study. *European Journal of Heart Failure* 2011; **14**: 333–340.
5. Njagi NE, Rizopoulos D, Molenberghs G, et al. A Joint Survival-Longitudinal Modelling Approach for the Dynamic Prediction of Rehospitalization in Telemonitored Chronic Heart Failure Patients. *Statistical Modelling* 2013; to appear.
6. Duchateau L and Janssen P. *The Frailty Model*. New York: Springer, 2008.
7. Tsiatis A and Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 2004; **14**: 809–834.
8. Verbeke G, Molenberghs G and Rizopoulos D. Random effects models for longitudinal data. *Longitudinal Research with Latent Variables*. K. van Montfort, H. Oud, and Al Satorra (Eds.). New York: Springer, 2010, pp. 37–96.
9. Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-event Data. *Biometrics* 2011; **67**: 819–829.
10. Molenberghs G, Verbeke G, et al. A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science* 2010; **25**: 325–347.
11. Molenberghs G, Verbeke G and Demétrio C. An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis* 2007; **13**: 513–531.
12. Molenberghs G and Verbeke G. *Models for Discrete Longitudinal Data*. New York: Springer, 2005, p.461.
13. Agresti A. *Categorical Data Analysis*. 2nd ed. New York: John Wiley & Sons, 2002, p.131.
14. Harrell FE Jr, Lee KL, Mark DB. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, **15**: 361–387.
15. Rizopoulos D. *Joint Models for Longitudinal and Time-to-event Data*. Boca Raton: Chapman and Hall/CRC, 2012, p. 171.
16. Rizopoulos D. Package “JM”: R package version 1.0-0, <http://cran.r-project.org/web/packages/JM/JM.pdf> (2012, accessed 16 February 2013).