

Robust Frailty Modelling using Non-Proportional Hazards Models

Il Do Ha¹ and Gilbert MacKenzie²

¹ Department of Asset Management, Daegu Haany University, Gyeongsan, 712-715, South Korea Email: idha@dhu.ac.kr

² Centre of Biostatistics, Department of Mathematics & Statistics, University of Limerick, Limerick, Ireland & ENSAI, Rennes, France. Email: gilbert.mackenzie@ul.ie

Abstract. Correlated survival times can be modelled by introducing a random effect, or frailty component, into the hazard function. For multivariate survival data we extend a non-PH model, the generalized time-dependent logistic survival model, to include random effects. The hierarchical-likelihood procedure, which obviates the need for marginalization over the random effect distribution, is derived for this extended model and its properties discussed. The extended model leads to a robust estimation result for the regression parameters against the mis-specification of the form of the basic hazard function or frailty distribution compared to PH-based alternatives. The proposed method is illustrated by two practical examples and a simulation study which demonstrate the advantages of the new model.

Keywords: Frailty models; Generalized time-dependent logistic (GTDL); hierarchical likelihood; Non-PH model; Random effect

²Correspondence to: Prof. G. MacKenzie, Centre of Biostatistics, University of Limerick, Ireland.

1. Introduction

Multivariate (or correlated) survival data are frequently encountered in biomedical researches where clustered, or multiple event, times are observed. The correlation between survival times has been usually modelled by introducing a frailty component (random effect) into the hazard function. These generalisations of Cox's (1972) semi-parametric model have now been widely used. In these models the conditional hazard function given the frailty is assumed to follow a standard proportional hazards (PH) model. The resulting marginal model, obtained by integrating out the frailty term is generally non-PH, except when the frailty follows a positive stable distribution (Hougaard, 2000, pp. 232; Hsu, Gorfine and Malone, 2007). Exactly how the marginal model deviates from proportionality is not known for many frailty distributions, including lognormal distribution (Hougaard, 2000, pp. 245).

As an alternative to the PH model, MacKenzie (1996) introduced the generalized time-dependent logistic (GTDL) non-PH model for univariate survival data. This model is a wholly parametric competitor for the PH model which generalizes the relative risk (RR) in Cox's semiparametric PH model to time-dependent form. It is thus able to accommodate a wider class of univariate survival data including PH survival data. Accordingly, in this paper we introduce a flexible non-PH frailty model for multivariate survival data based on the GTDL model.

For inference we use hierarchical likelihood (h-likelihood, Lee and Nelder, 1996, 2001) which obviates the need for marginalization over the frailty distribution. The h-likelihood approach provides a unified inferential framework and a numerically efficient fitting algorithm for various random-effect models including frailty models (Ha et al., 2001; Lee, Nelder and Pawitan, 2006).

In general, frailty models require the specification of two main terms, the basic hazard function, which is usually multiplied by the frailty term, and the assumed frailty distribution. Here, the basic hazard depends on time and fixed covariates. It has been shown that parametric inference on the regression parameters can be sensitive to the choice of the hazard function. In parametric PH models, this amounts to sensitivity to different choices of the baseline hazard function (Ha and Lee, 2003). Moreover, the frailty is an unobservable variable and therefore it may not always be easy to

check the goodness of fit of the frailty model (Ferreira and Garcia, 2001; Hsu et al., 2007). Accordingly, here, we study the effect of the mis-specification of the basic hazard function and the frailty distribution on the regression parameter estimates in various scenarios. We also compare the GTDL frailty model and Cox's PH frailty model, since both models are non-PH. The simulation studies show that the proposed GTDL frailty model is comparable to Cox's PH frailty model and inference on the regression parameter is robust against mis-specification of the basic hazard function and/or the frailty distribution.

The paper is organized as follows. In Section 2 we review the GTDL model briefly, while in Section 3 we formulate the extended model based on the GTDL. The likelihood approach to inference is developed in Section 4, and the proposed method is illustrated using two well-known data sets in Section 5 and using simulation studies in Section 6, the results being compared with those obtained from the corresponding PH frailty models. Finally, some further discussion is given in Section 7.

2. The GTDL regression model

A non-PH model, the GTDL regression model (MacKenzie, 1996) is defined by the hazard function:

$$\lambda(t; x) = \lambda_0 p(t; x), \quad (1)$$

where $\lambda_0 > 0$ is a scalar, $p(t; x) = \exp(t\alpha + x^T\beta) / \{1 + \exp(t\alpha + x^T\beta)\}$ is a linear logistic function in time, α is a scalar measuring the effect of time and $\beta = (\beta_0, \beta_1, \dots, \beta_p)^T$ is a $(p + 1) \times 1$ vector of regression parameters associated with fixed covariates $x = (x_0, x_1, \dots, x_p)^T$ and $x_0 \equiv 1$. In this model the β s play the usual regression coefficient role and it may be shown that they measure the linear influence of the covariates on a generalised log odds scale, rather than on the log-linear scale, as in Cox's PH model. The time dependent relative risk, $RR(t)$, the ratio of hazard rates for two subjects with different covariate vectors, x_1 and x_2 , is given by

$$\rho(t; x_1, x_2) = \lambda(t; x_1) / \lambda(t; x_2) = \exp\{(x_1 - x_2)^T\beta\} \psi(t; x_1, x_2), \quad (2)$$

where

$$\psi(t; x_1, x_2) = \frac{\{1 + \exp(t\alpha + x_2^T\beta)\}}{\{1 + \exp(t\alpha + x_1^T\beta)\}}.$$

The leading term on the right hand side of (2) is Cox's constant of proportionality (the RR in a PH model) and thus in the GTDL model this constant is moderated by $\psi(\cdot)$, a function of both time and covariates, demonstrating, unequivocally, that the model is non-PH. Moreover, it should be noted that (2) does not depend on the parameter λ_0 . This model (1) also provides the hazards of various shapes, leading to a flexible survival regression model (MacKenzie, 1996). When $\alpha = 0$ the relative risk does not depend on time and from (1) and (2) the resulting model is PH - an exponential with $\lambda(t; x) = \lambda_0 p(x)$, i.e., a multiple of the usual multiple logistic function (Cox, 1970). In particular, it is often reasonable to assume that the relative risk is not constant with time but converges to unity eventually because in old age there is apparently no difference between treatments (i.e. the treatment effects may fade out gradually in the long run): see for example Hougaard (1991) and Royston and Parmar (2002). This is a common feature of non-PH models shared by the GTDL, but one which PH models such as the Weibull and Cox models do not possess.

The genesis and development of the GTDL are discussed in Blagojevic and MacKenzie (2007). The leading constant, λ_0 , in (1) was introduced originally to avoid a bounded hazard model. Later it was confirmed (MacKenzie, 2002) that λ_0 was not estimable directly by the method of maximum likelihood because λ_0 is aliased with the intercept term (β_0) in the regression predictor. This latter term is, however, required to ensure that the maximum likelihood estimators (MLEs) are invariant to the choice of reference category (Gillon and MacKenzie, 2004) in the design matrix parametrization of categorical variables. This lack of invariance, when x_0 is omitted in favour of λ_0 , the so-called Canonical TDL (CTDL) version of the model (MacKenzie, 2002), is only detectable in large sample sizes.

This has led, *inter alia*, to the view that the GTDL model is a building block for more general models via frailty. The GTDL-Gamma frailty model and extensions with structural dispersion have been presented recently (Blagojevic & MacKenzie, 2007; Lynch & MacKenzie, 2007, 2008). However, the TDL model (GTDL with $\lambda_0 = 1$) has been shown to fit a wide range of cancer and other survival data. Moreover, when, as here, λ_0 is replaced by the parametric frailty term u_i the (now unbounded) model

succeeds and when a prior distribution, $\pi(\lambda_0)$, exists for λ_0 , the (again unbounded) GTDL model also succeeds (Louzada-Neto et al, 2010). Thus, the GTDL model provides a basic non-PH comparator, which we extend, in this paper, to a log-Normal frailty setting.

Now, from (1), the cumulative hazard function is given explicitly by

$$\Lambda(t; x) = \int_0^t \lambda(s; x) ds = \frac{\lambda_0}{\alpha} \log \left\{ \frac{1 + \exp(t\alpha + x^T \beta)}{1 + \exp(x^T \beta)} \right\}. \quad (3)$$

Under non-informative censoring the ordinary censored-data likelihood, which depends on (1) and (3), is constructed and the MLEs for the parameters can be obtained using numerical methods such as Newton-Raphson.

For more details on the potential advantages of parametric competitors to Cox's model see MacKenzie (1996, 1997).

3. GTDL Frailty Model

The correlation between survival times, which arises in recurrent or multiple event times on the same subject, can be modelled by introducing a frailty, or random effect. We thus extend the non-PH model (1) to include frailty.

First we define the multivariate data structures as follows. Let T_{ij} ($i = 1, \dots, n$, $j = 1, \dots, n_i$, $N = \sum_i n_i$) be the survival time for the j th observation of the i th subject and C_{ij} be the corresponding censoring time. Let the observable random variables be $Y_{ij} = \min(T_{ij}, C_{ij})$ and $\delta_{ij} = I(T_{ij} \leq C_{ij})$, where $I(\cdot)$ is the indicator function. Denote by U_i the random variable denoting the unobserved frailty (or random effect) for the i th subject. We extend the model (1) to the multivariate survival data setting by inclusion of a frailty term acting multiplicatively on the individual hazard rate of (1). The GTDL non-PH frailty model is then defined as follows. Given $U_i = u_i$, the conditional hazard function of T_{ij} takes the form

$$\lambda_{ij}(t|u_i; x_{ij}) = \lambda_{ij}(t; x_{ij})u_i, \quad (4)$$

where $\lambda_{ij}(t; x_{ij})$ is a basic hazard function not depending on u_i and from (1) it is given by

$$\lambda_{ij}(t; x_{ij}) = \lambda_0 p(t; x_{ij})$$

with $p(t; x_{ij}) = [\exp(t\alpha + x_{ij}^T\beta) / \{1 + \exp(t\alpha + x_{ij}^T\beta)\}]$. Here $x_{ij} = (x_{ij0}, x_{ij1}, \dots, x_{ijp})^T$ with $x_{ij0} \equiv 1$. As explained above, hereafter we take $\lambda_0 = 1$ in (4). The frailties U_i are assumed to be independent and identically distributed random variables with a density function depending on the frailty parameter θ .

Recall that if U_i is log-normal or gamma, then $V_i = \log U_i$, the log-frailty, becomes Normal or log-Gamma, respectively. The corresponding model is usually called log-normal or gamma frailty model, based on frailty U_i . For convenience, we also shall refer to it as normal or log-gamma frailty model, based on log-frailty V_i .

If a basic hazard function in (4) is of the form

$$\lambda_{ij}(t; x_{ij}) = \lambda_0(t) \exp(x_{ij}^T\beta), \quad (5)$$

where $\lambda_0(t)$ is a baseline hazard function, we have a PH frailty model. Here, the term $x_{ij}^T\beta$ in (5) does not include an intercept term because of identifiable purposes. Note that $\lambda_0(t)$ can be parametric (e.g. Weibull) or non-parametric. In particular, the latter gives a semi-parametric Cox-PH frailty model, an extension of Cox's PH model (1972): see for example McGilchrist and Aisbett (1991) and Ha et al. (2001).

Remark 1: Following Hougaard (2000, pp. 226), the marginal hazard model, denoted by $\lambda^M(t; x)$, can be derived by integrating out the frailty from the conditional hazard function (4); it is thus given by

$$\lambda^M(t; x) = \lambda(t; x)E(U|T > t),$$

a convenient product. The conditional expectation, $E(U|T > t)$, can be calculated from the Laplace transform. For gamma frailty the computation is analytic, but for log-normal frailty numerical integration is required. As mentioned earlier, the marginal model, $\lambda^M(t; x)$, will not be PH unless U is positive stable. In particular, when U is log-normal or gamma frailty the non-PH model (4) and the PH model (5) are marginally non-PH. Accordingly, this affords an interesting opportunity to compare the behaviour of these two competing models.

4. Estimation procedure

Following Lee and Nelder (1996) and Ha et al. (2001), the h-likelihood for the model

(4), denoted by h , is defined by

$$h = h(\alpha, \beta, \theta) = \sum_{ij} \ell_{1ij} + \sum_i \ell_{2i}, \quad (6)$$

where

$$\begin{aligned} \ell_{1ij} &= \ell_{1ij}(\alpha, \beta; y_{ij}, \delta_{ij} | u_i) \\ &= \delta_{ij} \log \lambda_{ij}(y_{ij} | u_i) - \Lambda_{ij}(y_{ij} | u_i) \\ &= \delta_{ij} (\log p_{ij} + v_i) + u_i \alpha^{-1} \log(q_{ij} g_{ij}) \end{aligned}$$

is the logarithm of the conditional density function for Y_{ij} and δ_{ij} given $U_i = u_i$, which is also the ordinary censored-data log-likelihood given u_i , and $\ell_{2i} = \ell_{2i}(\theta; v_i)$ is the logarithm of the density function for $V_i = \log U_i$ with parameter θ . Here, the conditional hazard is $\lambda_{ij}(y_{ij} | u_i) = p_{ij} u_i$ where

$$p_{ij} = p_{ij}(\alpha, \beta) = \exp(y_{ij} \alpha + x_{ij}^T \beta) / \{1 + \exp(y_{ij} \alpha + x_{ij}^T \beta)\}$$

and the conditional cumulative hazard $\Lambda_{ij}(y_{ij} | u_i) = -u_i \alpha^{-1} \log(q_{ij} g_{ij})$ with $q_{ij} = 1 - p_{ij}$ and $g_{ij} = g_{ij}(\beta) = 1 + \exp(x_{ij}^T \beta)$.

Given frailty parameter θ , the maximum h-likelihood (MHL) joint estimating equations of $\tau = (\alpha, \beta^T, v^T)^T$ with $v = (v_1, \dots, v_n)^T$ are given by

$$\partial h / \partial \tau = 0,$$

leading to the detailed score equations:

$$\begin{aligned} \partial h / \partial \alpha &= \sum_{ij} \{ \delta_{ij} q_{ij} y_{ij} - (u_i / \alpha) p_{ij} y_{ij} - (u_i / \alpha^2) \log(q_{ij} g_{ij}) \}, \\ \partial h / \partial \beta_k &= \sum_{ij} \{ \delta_{ij} q_{ij} + (u_i / \alpha) (r_{ij} - p_{ij}) \} x_{ijk} \quad (k = 0, 1, \dots, p), \\ \partial h / \partial v_i &= \sum_j \{ \delta_{ij} + (u_i / \alpha) \log(q_{ij} g_{ij}) \} + \partial \ell_{2i} / \partial v_i \quad (i = 1, \dots, n), \end{aligned} \quad (7)$$

where $u_i = \exp(v_i)$ and $r_{ij} = r_{ij}(\beta) = \exp(x_{ij}^T \beta) / \{1 + \exp(x_{ij}^T \beta)\}$. The estimating equations (7) are easily solved using the Newton-Raphson method. The asymptotic covariance matrix for $\hat{\tau} - \tau$ is given by the inverse of $H = H(h, \tau, \theta) = - \partial^2 h / \partial \tau^2$

(Lee and Nelder, 1996; Ha et al., 2001). Notice here that H depends on the fixed parameters, (α, β, θ) .

For inference of θ , we use the adjusted profile h-likelihood (Lee and Nelder, 2001), defined by

$$p_\tau(h) = [h - \frac{1}{2} \log \det\{H(h, \tau, \theta)/(2\pi)\}]|_{\tau=\hat{\tau}} \quad (8)$$

where $\hat{\tau} = \hat{\tau}(\theta) = (\hat{\alpha}(\theta), \hat{\beta}^T(\theta), \hat{v}^T(\theta))^T$. Note here that $p_\tau(h)$ is a function of θ only because it has already eliminated τ from h . The REML (restricted maximum likelihood) estimating equation for θ , maximizing $p_\tau(h)$ of (8), is given by

$$\partial p_\tau(h)/\partial \theta = 0. \quad (9)$$

In model (4) with log-normal frailty, where $V_i = \log U_i \sim N(0, \theta)$, the equation (9) gives the REML estimator

$$\hat{\theta} = (\hat{v}^T \hat{v})/(n - \gamma), \quad (10)$$

where $\gamma = \text{trace}(K)/\theta$ and K is the matrix given by the bottom right-hand corner of H^{-1} (McGilchrist, 1993; Ha et al., 2001). Furthermore, for model (4) with the gamma frailty having $E(U_i) = 1$ and $\text{var}(U_i) = \theta$, we use the second-order Laplace approximation (Lee and Nelder, 2001; Ha and Lee, 2003, 2005).

Remark 2: Marginal likelihood, denoted by m , is an alternative vehicle for inference; it can be obtained by integrating over the frailty in the h-likelihood:

$$m = m(\alpha, \beta, \theta) = \sum_i \log \left\{ \int \exp(h_i) dv_i \right\}, \quad (11)$$

where $h_i = \sum_j \ell_{1ij} + \ell_{2i}$ is the contribution of the i th individual to h in (6). For model (4) with gamma frailty having $E(U_i) = 1$ and $\text{var}(U_i) = \theta$, we have from (11) an explicit marginal likelihood m as in PH gamma frailty models:

$$\sum_{ij} \delta_{ij} \log p_{ij} + \sum_i \{ -(\theta^{-1} + \delta_{i+}) \log(\theta^{-1} + \Lambda_{i+}) + \log \Gamma(\theta^{-1} + \delta_{i+}) - c(\theta) \},$$

where $\delta_{i+} = \sum_j \delta_{ij}$, $\Lambda_{i+} = \sum_j \Lambda(y_{ij}) = -\sum_j \alpha^{-1} \log(q_{ij}g_{ij})$ and $c(\theta) = -\log \Gamma(\theta^{-1}) - \theta^{-1} \log \theta$. Ha et al. (2001) showed that in Cox's PH-gamma frailty models, given θ , the MHL estimator of β is the same as the marginal maximum likelihood estimator. Similarly, we can show that in GTDL-gamma frailty models this fact still holds for α

and β . However, in general, the marginal likelihood does not lead to a closed form, whence complicated numerical integration methods may be required.

5. Examples

We illustrate the use of the proposed model by analyzing two well-known multivariate survival data sets. For the purposes of comparison, we include Cox’s PH frailty model. Even though the results of this paper can be applied to non-normal distributions (e.g. log-gamma) for log-frailty V_i , for simplicity of argument, we employ a normal distribution for V_i , which gives a simple estimating equation (10) for frailty parameter θ under h-likelihood framework (Ha et al., 2001; Ha and Lee, 2003). In particular, the normal assumption for V_i is very useful for modelling multi-component or correlated frailties (Yau, 2001; Ha, Lee and MacKenzie, 2007). Accordingly, for the comparison we fit two log-normal frailty models: the GTDL non-PH frailty model (GFM) and Cox’s PH frailty model (CFM).

5.1. Litter-matched rat data

Mantel et al. (1977) presented a data set on a tumorigenesis study of 50 litters of female rats. For each litter, one rat was selected to receive the drug and the other two rats were placebo-treated controls. The survival time is the time to development of tumor, measured in weeks. Death before occurrence of tumor yields a right-censored observation; forty rats developed a tumor, leading to 73% censoring. The survival times for rats in a given litter may be correlated due to a random effect representing shared genetic or environmental effects. The two frailty models were fitted with a single fixed covariate (x_{ij}) having $x_{ij}^T \beta = \beta_0 + \beta_1 x_{ij}$ ($i = 1, \dots, 50; j = 1, 2, 3$). Here $x_{ij} = 1$ if the j th rat in the i th litter received the drug and 0 otherwise.

The results of fitting the two models (GFM & CFM) are summarized in Table 1. In GFM the estimated time coefficient $\hat{\alpha} = 0.048$ (with SE=0.008) suggests that the effect of time is significantly different from zero, indicating an increasing time-trend in hazard. The intercept term is also well-defined. In addition, both models give similar estimates of the treatment effect and frailty parameter, (β_1, σ) where $\sigma = \sqrt{\theta}$ is the square root of the log-frailty variance in $N(0, \theta)$. In particular, the Wald test of

$H_0 : \beta_1 = 0$, yields similar χ^2 -values (p -values) with 1 d.f., namely: 7.94 (0.0048) for GFM and 7.86 (0.0050) for CFM, respectively.

Table 1 near here

5.2. CGD data

Fleming and Harrington (1991) provided more extensive multivariate survival data on a placebo-controlled randomized trial of gamma interferon (γ -IFN) in chronic granulomatous disease (CGD). The aim of the trial was to investigate the effectiveness of the γ -IFN in reducing the rate of serious infections in CGD patients. In this study, 128 patients were followed for approximately 1 year. Out of the 63 patients in the treatment group, 14 patients experienced at least one infection and a total of 20 infections were recorded. In the placebo group, 30 out of 65 patients experienced at least one infection, with a total of 56 infections being recorded. Here, the survival times are the times between recurrent CGD infections on each patient (i.e., gap times). Censoring occurred at the last observation on all patients, except one, who experienced a serious infection on the date he left the study. The recurrent infection times for each patient are likely to be correlated as in the litter-matched rat data study.

We fitted the same set of fixed covariates considered by Yau and McGilchrist (1998), namely: treatment (0=placebo, 1= γ -IFN), pattern of inheritance (0=autosomal recessive, 1=X-linked); age (in years); height (in cm); weight (in kg); using corticosteroids at time of study entry (0=no, 1=yes); using prophylactic antibiotics at time of study entry (0=no, 1=yes); sex (0=male, 1=female), hospital region (0=U.S., 1=Europe), and a longitudinal variable representing the accumulated time from the first infection in years. The rationale under-pinning this variable is that the infection rate may increase over time following the first infection (Yau and McGilchrist, 1998).

Table 2 near here

The results pertaining to the two models are given in Table 2. Overall, both results are very similar as in the rat data of Section 5.1, except for the findings in relation to time-dependent effects. In particular, the interpretation of the beneficial effect of treatment is unequivocal in both models. The time parameter α is statistically significant, suggesting that the non-PH model (GFM) may be appropriate. However, the

magnitude of the dependence of the hazard on time is so small as to be almost immaterial in this case. However, the GFM suggests the existence of a longitudinal effect, while the CFM does not it. And, further evidence for its existence, may be adduced from the analysis by Yau and McGilchrist (1998) which identified a longitudinal effect using a more complicated PH model with time-dependent AR(1) frailties.

Remark 3 Overall, both models yield similar results in the data sets studied. In some ways this is perhaps not too surprising since both frailty models are non-PH. However, there is some disagreement about the significance of the longitudinal effect in the CDG data set. The ordinary interpretation is that we should at least suspect the existence of such an effect when the models disagree and, in this case, regard the illuminating analysis by Yau and McGilchrist (1998) as confirmatory. Our approach highlights the advantages of fitting alternative models such as the flexible, non-PH, GTDL to the same dataset, rather than drawing conclusions from a single model class which in practice is typically just PH.

However, it remains to be seen how, if at all, the two frailty models differ and to get explore this issue we describe a detailed simulation study in the next section.

6. Simulation Study

Numerical studies, based upon 500 replications of simulated data, are presented to evaluate the performance and robustness of the proposed model compared to competing frailty models.

6.1. Simulation scheme

We based the simulation strategy on the structure of litter-match rat data in Section 5.1. Data for analysis were generated from the following frailty models.

$$\lambda_{ij}(t|v_i; x_{ij}) = \lambda_{ij}(t; \eta_{ij}) \exp(v_i), \quad (12)$$

where $\eta_{ij} = \beta_0 + \beta_1 x_{ij}$, $v_i = \log u_i$, and $i = 1, \dots, 50$ and $j = 1, 2, 3$. Here, the covariate values are generated by two different designs. Under the first design, $x_{i1} = 1$ and $x_{i2} = x_{i3} = 0$ for $i = 1, \dots, 50$. Under the second design, $x_{i1} = x_{i2} = x_{i3} = 1$ for the first 25 subjects (the treatment group) and 0 for the remaining 25 (the control

group). Notice here that the first design corresponds to the same group-matched study as analyzed in Table 1 and the second design to the group randomization study (in which the group is the randomization unit) as in the treatment covariate of Table 2. The corresponding censoring times were generated from uniform distribution with the parameter values determined empirically to achieve approximately two right censoring rates, low (around 20%) and high (around 73%).

As in Section 5 we assumed a normal distribution for log-frailty v . We also used the following combinations for the basic hazard $\lambda(t; \eta)$ and log-frailty v in the simulation model (12) to investigate the robustness of the proposed method against violations of model assumptions on (a) the basic hazard function and (b) the frailty distribution:

- (i) GTDL hazard (G): $\lambda(t; \eta) = \exp(\alpha t + \eta) / \{1 + \exp(\alpha t + \eta)\}$,
- (ii) Weibull hazard (W): $\lambda(t; \eta) = \phi t^{\phi-1} \exp(\eta)$,
- (iii) Normal log-frailty (N): $v \sim N(0, \sigma^2)$
- (iv) Extreme value log-frailty (EV): $v \sim EV(0, \sigma^2)$.

Thus, in this simulation, we are interested in the effect of mis-specification of these two components.

Here, $\eta = \beta_0 + \beta_1 x$ and $EV(0, \sigma^2)$ denotes an extreme value distribution with a shift such that its mean is 0 and variance σ^2 . For the values of parameters used in the simulation we employ the actual estimates from the GTDL frailty model in Table 1; $\alpha = 0.048$, $\beta_0 = -9.271$, $\beta_1 = 0.927$ and $\sigma = 0.679$. For Weibull shape parameter in (ii) we set $\phi = 1.5$. For the basic hazard (i) is non-PH, while (ii) is PH. The frailty defined by (iii) is a symmetric distribution about mean 0, whereas that defined by (iv) is a skewed distribution (Ha, Lee and Song, 2002).

In model (12), when the form of $\lambda(t; \eta)$ is GTDL and the distribution of v is normal, we designate this combination as the G-N model. Thus, we simulate from the following four models: G-N, G-EV, W-N and W-EV. For each of the four simulation models, 500 simulated data sets were generated. We present a subset of our simulation results in which three lognormal frailty models (G-N, C-N and W-N) were fitted to all of the data sets using the h-likelihood methods described above. Here C-N indicates Cox-PH model with normal log-frailty.

For the 500 replications we computed the percentage of relative bias (RB%), the standard deviation (SD), the mean of the estimated standard error (SE) and the mean squared error (MSE) for $\hat{\beta}_1$. The RB% for $\hat{\beta}_1$ is defined by

$$\text{RB\%} = \left\{ \frac{\bar{\beta}_1 - \beta_1}{\beta_1} \right\} \times 100,$$

where $\bar{\beta}_1 = \sum_i \hat{\beta}_1^{(i)} / 500$ is the mean of the $\hat{\beta}_1^{(i)}$'s and $\hat{\beta}_1^{(i)}$ is the estimate of β_1 in the i th replication, and the SD and MSE are, respectively, defined as

$$\text{SD} = \left\{ \sum_i (\hat{\beta}_1^{(i)} - \bar{\beta}_1)^2 / 499 \right\}^{1/2} \quad \text{and} \quad \text{MSE} = \sum_i (\hat{\beta}_1^{(i)} - \beta_1)^2 / 500.$$

The SE for $\hat{\beta}_1$ is obtained from H^{-1} . In addition, we calculated the empirical coverage probability for a nominal 95% confidence interval for β_1 based on the SE. For the frailty parameter σ the corresponding RB%, SD and MSE are also given. For the computation we used SAS/IML.

6.2. Simulation results

The results for the 20% censoring rate are summarized in Table 3. We report the estimation results for the parameters of interest, (β_1, σ) , in the fitted models.

Tables 3 and 4 near here

Firstly, the proposed model (G-N) overall performs well in all cases considered. Under design 1, as expected, fitting the G-N model shows larger biases when basic hazard and/or frailty distribution are misspecified (i.e. under G-EV, W-N and W-EV) than when they are correctly specified (i.e. under G-N). In particular, fitting the G-N leads to a RB% of -2.7 in $\hat{\sigma}$ when both basic hazard and frailty distribution are misspecified (i.e. under W-EV), but the relative bias of β_1 is relatively small as in RB% = -1.5. The coverage probabilities for β_1 are also reasonable with the 93.8%–95.0% range under the four true models considered. And our standard-error estimates also work well as judged by the very good agreement between SE and SD.

It should be noticed that when the true model is G-N, the SD and SE entries for $\hat{\beta}_1$ tend to be smaller under the alternative mis-specified models. This arises, in part, from the fact that the G-N model fits more parameters than the competing Cox's PH frailty

model (C-N) and so the generalized variance is typically larger; see also simulation results by Kuk and Chen (1992). Under this situation fitting the misspecified Weibull frailty model (W-N) also shows the smaller SD and SE, which are partly caused by the fact that fitting the W-N gives seriously downward biased estimates in both β_1 and σ . These mean that when under the true G-N model a competing model is used the standard errors will be optimistic.

Under design 2, as expected, the variation in SD, SE and MSE, for $\hat{\beta}_1$ are higher than in design 1. However, the G-N still performs as well as in design 1. Again these results suggest that the proposed model is satisfactory when the regression parameters are the subject of the inference.

Secondly, we find that under both designs the trends from fitting the Cox's PH frailty model (C-N) are similar to those evident in G-N, even if the G-N is sometimes less biased. In particular, the robustness of the C-N model in relation to β_1 confirms the simulation results of Pickles and Crouchley (1995), Ng and Cook (2000) and Ha and Lee (2003, 2005).

Finally, when fitting Weibull-PH frailty model (W-N) it performs well for the estimation of β_1 if the basic hazard is correctly specified (i.e. under W-N & W-EV). However, it shows seriously downward biases for the estimation of both β_1 and σ when the basic hazard is misspecified (i.e. under G-N & G-EV). In particular, under design 2 the coverage probabilities for β_1 from fitting W-N are noticeably underestimated, with 92.0% under the true G-N model and 91.0% under the true G-EV model.

Table 4 shows the results for the high censoring simulation (73%). Overall, the trends are similar overall to those observed in Table 3.

Table 5 near here

Furthermore, under the setting in Table 4 the α and σ were increased to $(\alpha, \sigma) = (0.15, 1.0)$, implying a stronger non-PH effect and a higher correlation among the survival times. The corresponding simulation results are given in Table 5. Overall, the trends in the findings are similar to those of Tables 3 and 4. However, one rather striking difference is the smaller standard errors caused by the under-estimation of (β_1, σ) in the competing models when the G-N model is true. The under-estimation of the uncertainty associated with β_1 is relatively large for the C-N model, but is more serious

in the W-N model. In particular, under the true G-N model in design 1 the coverage probabilities for β_1 from fitting C-N are underestimated, at 92.2%. Thus, we have identified a penalty arising from mis-specification of the non-PH basic hazard function in the stronger non-PH and larger frailty variance cases. It should be noted that the G-N model is immune to this effect.

In summary, we conclude the proposed model gives a robust result for the estimation of β_1 even if the model assumptions about baseline hazard and/or frailty distribution are violated, and it is a fully parametric competitor for the Cox-PH frailty model considered here.

7. Discussion

Our findings indicate that when modelling survival data it is not necessary to start from an initial PH assumption for the basic hazard. Indeed in an increasing number of these situations the usual PH assumption is found to be untenable and consequently there is a need for alternative models.

This consideration led us to a new non-PH model based on the GTDL family (MacKenzie, 1996, 1997). The GTDL frailty model was easily implemented in the h-likelihood framework, and has proved a flexible tool for analyzing correlated data. We have found via the two examples and simulation studies that the proposed model may be viewed as a wholly parametric competitor for Cox's PH frailty model.

We also note that inference for Cox frailty models is usually complicated because of issues surrounding the infinite-dimensional nuisance parameters which may arise in the estimation of the baseline hazard function, thereby violating the usual regular estimating assumptions (Ha and Lee, 2005). By contrast, the GTDL frailty model is free from this particular complication and being wholly parametric all of the time-dependent quantities of interest (e.g. hazard trend) may be readily derived (Royston and Parmar, 2002).

In the two example data sets analyzed we observe that the time effect parameters (α) are both significant, but their magnitudes are relatively small. Accordingly, both models (i.e. G-N and C-N) give similar estimates of regression parameters, a finding confirmed by the simulation results of Tables 3 and 4 (with $\alpha = 0.048$). However, in

the analysis of CGD data, the GTDL frailty model signalled the need for a longitudinal effect and this was confirmed in further, independent, analysis.

However, the simulation results (Table 5) show that with a larger time-effect parameter both models give different results for β_1 and that G-N model is less biased. More worrying for the PH class, particularly the W-N, is the finding that its standard errors are artificially precise. This finding alone justifies our wider approach to data analysis where we fit a variety of models from different classes to the data.

The simulation results also show that the proposed model works well for regression parameters β_1 , even when basic hazard function or frailty distribution is mis-specified. This suggests that the GTDL frailty model can be a practical choice in real data analysis if the regression parameters are of primary interest, as in multi-centre clinical trial with survival data (Gray, 1994; Vaida and Xu, 2000) which can lead a heterogeneity between centres, when the choice of underlying basic hazard or frailty distribution is not straightforward.

However, the simulation shows that none of the models are particularly successful at estimating the frailty parameter σ , when the frailty distribution is misspecified; see for example Ferreira and Garcia (2001), Agresti et al. (2004) and Ha and Lee (2005). This problem may be overcome by using heavy-tailed distributions such as t-distribution (Noh et al., 2005 and Lee and Nelder, 2006) or a nonparametric distribution (Laird, 1978; Ng and Cook, 2000).

Accordingly, we conclude that there is clearly considerable scope for expanding the class of non-PH models available in this setting and our methods can be extended to GTDL models with multi-component frailties (Ha et al., 2007) or correlated frailties (Yau and McGilchrist, 1998; Ripatti and Palmgren, 2000). And finally we note that extending the GTDL frailty model to incorporate more robust frailty distributions will be the subject of a future communication.

Acknowledgments

This work was supported by the Korea Research Foundation Grant (KRF-2006-013-C00092) and Science Foundation Ireland (BIO-SI, project, 07/MI/012 see www.ul.ie/bio-si). The work was carried out while the first author was visiting the Centre of Biostatistics in University of Limerick, Ireland (2006-7). The paper was finalized for publication when the second author was Visiting Professor of Statistics, ENSAI, Rennes France (January-July, 2010). We particularly wish to thank two anonymous referees for their perceptive comments which improved the paper.

References

- Agresti, A., Caffo, B. and Ohman-Strickland, P. (2004). Examples in which misspecification of a random effect distribution reduces efficiency, and possible remedies. *Computational Statistics and Data Analysis*, **47**, 639-653.
- Blagojevic, M. & MacKenzie, G. (2007). A comparison of PH and non-PH frailty models. In: *Correlated Data Modelling 2004*, 239-249. Editors: Dario Gregori, Gilbert MacKenzie, Herwig Friedl, Roberto Corradetti. Franco Angeli, Milano, Italy.
- Cox, D. R. (1970) *The Analysis of Binary Data*. London: Methuen.
- Cox, D. R. (1972) Regression models and life-tables (with discussion). *J. R. Statist. Soc. B*, **34**, 187-220.
- Ferreira, A. and Garcia, N. L. (2001) Simulation study for misspecifications on a frailty model. *Brazilian Journal of Probability and Statistics*, **15**, 121-134.
- Fleming, T. R. and Harrington, D. P. (1991) *Counting processes and survival analysis*. New York: Wiley.
- Gillon, N. and MacKenzie, G (2004) The North Staffordshire Breast Cancer Study. Centre for Medical Statistics, Keele University, England: Internal Technical Report Series. No 2, pp.1-80.
- Gray, R.J. (1994) A Bayesian analysis of institutional effects in multicenter cancer clinical trial. *Biometrics*, **50**, 244-253.

- Ha, I. D. and Lee, Y. (2003) Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*, **12**, 663-681.
- Ha, I. D. and Lee, Y. (2005) Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. *Biometrika*, **92**, 717-723.
- Ha, I. D., Lee, Y. and MacKenzie, G. (2007) Model selection for multi-component frailty models. *Statistics in Medicine*, **26**, 4790-4807.
- Ha, I. D., Lee, Y. and Song, J.-K. (2001) Hierarchical likelihood approach for frailty models. *Biometrika*, **88**, 233-243.
- Ha, I. D., Lee, Y. and Song, J.-K. (2002) Hierarchical likelihood approach for mixed linear models with censored data. *Lifetime Data Analysis*, **8**, 163-176.
- Hougaard, P. (1991) Modelling heterogeneity in survival data. *Journal of Applied Probability*, **28**, 695-701.
- Hougaard, P. (2000) *Analysis of multivariate survival Data*. New York: Springer-Verlag.
- Hsu, L., Gorfine, M. and Malone, K. (2007) On robustness of marginal regression coefficient estimates and hazard functions in multivariate survival analysis of family data when the frailty distribution is mis-specified. *Statistics in Medicine*, **26**, 4657-4678.
- Kuk, A. Y. C. and Chen, C.-H. (1992) A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, **79**, 531-541.
- Laird, N. (1978) Non-parametric maximum likelihood estimation of mixing distribution. *J. Am. Statist. Assoc.*, **73**, 805-811.
- Lee, Y. and Nelder, J. A. (1996) Hierarchical generalized linear models (with discussion). *J. R. Statist. Soc. B*, **58**, 619-678.
- Lee, Y. and Nelder, J. A. (2001) Hierarchical generalised linear models: a synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika*, **88**, 987-1006.
- Lee, Y. and Nelder, J. A. (2006) Double hierarchical generalised linear models (with discussion). *J. R. Statist. Soc. C*, **55**, 139-185.

- Lee, Y. and Nelder, J. A. and Pawitan, Y. (2006) *Generalised linear models with random effects: unified analysis via h-likelihood*. London: Chapman and Hall.
- Louzada-Neto F, Cremasco C.P. and MacKenzie, G. (2010) Bayesian Modelling for the GTDL survival model. *JSTA* vol.9, No 2 (June, in press).
- Lynch, J. and MacKenzie, G. (2007) The analysis of breast cancer survival in local health authorities. Proceedings of the 22nd International Workshop on Statistical Modelling (IWSM). Barcelona, Spain, July 2007, pp 408-411.
- Lynch, J. and MacKenzie, G. (2008) Non-PH survival modelling with frailty and structured dispersion. Contributed Oral Presentation, 24th International Biometric Conference, Dublin, Ireland July. Abstract may be found at:
- and presentation at www.ul.ie/biostatistics.
- Mantel, N., Bohidar, N. R. and Ciminera, J. L. (1977) Mantel-Haenszel analysis of litter-matched time-to-response data, with modifications for recovery of interlitter information. *Cancer Research* **37**, 3863-3868, 1977.
- MacKenzie, G. (1996) Regression models for survival data: the generalised time dependent logistic family. *J. R. Statist. Soc. D*, **45**, 21-34.
- MacKenzie, G. (1997) On a non-proportional hazards regression model for repeated medical random counts. *Statistics in Medicine*, **16**, 1831-1843.
- MacKenzie, G. (2002) A logistic regression model for survival data. Proceedings of the 17th International Workshop on Statistical Modelling (IWSM). Chania, Crete, July 2002, pp 431-438.
- McGilchrist, C. A. (1993) REML estimation for survival models with frailty. *Biometrics*, **49**, 221-225.
- McGilchrist, C. A. and Aisbett, C. W. (1991) Regression with frailty in survival analysis. *Biometrics*, **47**, 461-466.
- Ng, E. T. M. and Cook, R. J. (2000) A comparison of some random effect models for parameter estimation in recurrent events. *Mathematical and Computer Modelling*, **32**, 11-26.

- Noh, M, Lee, Y. and Pawitan, Y. (2005) Robust ascertainment-adjusted parameter estimation. *Genetic Epidemiology*, **29**, 68-75.
- Pickles, A. and Crouchley, R. (1995). A comparison of frailty models for multivariate survival data. *Statistics in Medicine*, **14**, 1447-1461.
- Ripatti, S. and Palmgren, J. (2000) Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, **56**, 1016-1022.
- Royston, P. and Parmar, M. K. B. (2002) Flexible parametric proportional-hazards and proportional-odds models for censored data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine*, **21**, 2175-2197.
- Yau, K. K. W. (2001) Multilevel models for survival analysis with random effects. *Biometrics* **57**, 96-102.
- Yau, K. K. W. and McGilchrist, C. A. (1998) ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine* **17**, 1201-1213.
- Vaida, F. and Xu, R. (2000) Proportional hazards models with random effects. *Statistics in Medicine*, **19**, 3309-3324.

NB: Working papers on the GTDL model will be available at <http://www.ul.ie/biostatistics> in October 2010.

Table 1. Results of fitting the two frailty models to the litter-matched rat data

Variable	GFM		CFM	
	Est.	SE	Est.	SE
Intercept	-9.271	0.645	—	—
Drug	0.927	0.327	0.903	0.322
Time (α)	0.048	0.008	—	—
Frailty (σ)	0.679	—	0.636	—

GFM, GTDL non-PH lognormal frailty model;
CFM, Cox's PH lognormal frailty model;
Est., estimate; SE, standard error;
 α , time effect of GTDL;
 σ , a square root of log-frailty variance in $N(0, \sigma^2)$

Table 2. Results of fitting the two frailty models to the CGD data

Variable	GFM		CFM	
	Est.	SE	Est.	SE
Intercept	-5.706	1.298	—	—
γ -IFN	-1.087	0.320	-1.105	0.324
Inheritance	-0.684	0.359	-0.650	0.358
Age	-0.087	0.043	-0.085	0.043
Height	0.008	0.013	0.008	0.013
Weight	0.011	0.020	0.010	0.020
Corticosteroids	2.039	0.801	1.962	0.795
Prophylactic	-0.660	0.422	-0.662	0.421
Sex	-0.796	0.498	-0.751	0.496
Hospital region	-0.712	0.377	-0.688	0.377
Longitudinal	1.062	0.500	0.914	0.505
Time (α)	0.0039	0.0015	—	—
Frailty (σ)	0.715	—	0.713	—

GFM, GTDL non-PH lognormal frailty model;
CFM, Cox's PH lognormal frailty model;
Est., estimate; SE, standard error;
 α , time effect of GTDL;
 σ , a square root of log-frailty variance in $N(0, \sigma^2)$

Table 3. The true values $(\alpha, \sigma) = (0.048, 0.679)$ and 20% censoring: simulation results of fitting three frailty models when basic hazard or log-frailty distribution is correctly specified or misspecified. The α and σ are, respectively, a time effect in GTDL model and a square root of log-frailty variance in $N(0, \sigma^2)$. The true regression and Weibull shape parameters are, respectively, $\beta_1 = 0.927$ and $\phi = 1.5$.

Design	True	Fitted	$\hat{\beta}_1$					$\hat{\sigma}$		
			RB%	SD	(SE)	MSE	95%	RB%	SD	MSE
1	G-N	G-N	0.2	0.231	(0.234)	0.053	0.948	1.2	0.209	0.043
		C-N	-1.6	0.233	(0.224)	0.055	0.946	1.5	0.220	0.048
		W-N	-12.3	0.193	(0.208)	0.050	0.932	-30.6	0.243	0.102
	G-EV	G-N	3.0	0.237	(0.223)	0.057	0.944	1.6	0.212	0.045
		C-N	-0.9	0.245	(0.224)	0.060	0.940	-1.3	0.213	0.045
		W-N	-9.9	0.214	(0.210)	0.054	0.936	-23.1	0.237	0.081
	W-N	G-N	-2.1	0.230	(0.218)	0.053	0.950	-1.9	0.189	0.036
		C-N	-0.1	0.229	(0.216)	0.052	0.944	-2.9	0.195	0.039
		W-N	0.1	0.229	(0.213)	0.052	0.950	-0.7	0.204	0.041
	W-EV	G-N	-1.5	0.220	(0.217)	0.049	0.938	-2.7	0.185	0.035
		C-N	-0.3	0.218	(0.216)	0.048	0.954	-2.0	0.175	0.031
		W-N	1.4	0.219	(0.214)	0.048	0.954	1.5	0.186	0.035
2	G-N	G-N	0.2	0.296	(0.298)	0.088	0.946	1.0	0.198	0.039
		C-N	-3.7	0.299	(0.286)	0.091	0.936	-0.4	0.212	0.045
		W-N	-13.2	0.242	(0.231)	0.074	0.920	-33.0	0.241	0.108
	G-EV	G-N	2.9	0.334	(0.311)	0.112	0.940	1.4	0.209	0.044
		C-N	-2.0	0.326	(0.315)	0.107	0.938	0.1	0.207	0.043
		W-N	-10.3	0.264	(0.240)	0.079	0.910	-24.4	0.250	0.090
	W-N	G-N	-0.4	0.283	(0.292)	0.080	0.952	-1.2	0.200	0.040
		C-N	1.0	0.281	(0.281)	0.079	0.948	-0.9	0.198	0.039
		W-N	0.3	0.281	(0.282)	0.079	0.952	1.5	0.202	0.041
	W-EV	G-N	-1.1	0.282	(0.274)	0.080	0.940	-2.5	0.192	0.037
		C-N	-0.4	0.279	(0.278)	0.078	0.944	-2.9	0.189	0.036
		W-N	0.8	0.279	(0.280)	0.078	0.946	0.9	0.209	0.044

The simulation is conducted with 500 replications using the structure of rat data ($N = 150$ with $n = 50$ & $n_i = 3$, and 20% censoring) in Section 5.1. In design 1, $x_{i1} = 1$ and $x_{i2} = x_{i3} = 0$ for $i = 1, \dots, 50$ and in design 2, $x_{i1} = x_{i2} = x_{i3} = 1$ for the first 25 subjects and 0 for the remaining 25. G-N, C-N, W-N, G-EV and W-EV denote GTDL-normal, Cox's PH-normal, Weibull-normal, GTDL-extreme value and Weibull-extreme value frailty models, respectively. RB% and 95%, respectively, indicate the percentage of relative bias and empirical coverage probability for a nominal 95% confidence interval for β_1 .

Table 4. The true values $(\alpha, \sigma) = (0.048, 0.679)$ and 73% censoring: simulation results of fitting three frailty models when basic hazard or log-frailty distribution is correctly specified or misspecified. The α and σ are, respectively, a time effect in GTDL model and a square root of log-frailty variance in $N(0, \sigma^2)$. The true regression and Weibull shape parameters are, respectively, $\beta_1 = 0.927$ and $\phi = 1.5$.

Design	True	Fitted	$\hat{\beta}_1$					$\hat{\sigma}$		
			RB%	SD	(SE)	MSE	95%	RB%	SD	MSE
1	G-N	G-N	0.3	0.393	(0.375)	0.154	0.946	-1.6	0.331	0.110
		C-N	-1.7	0.375	(0.365)	0.141	0.942	-8.1	0.329	0.111
		W-N	-14.1	0.325	(0.341)	0.123	0.936	-44.0	0.329	0.198
	G-EV	G-N	4.3	0.407	(0.373)	0.168	0.938	-6.8	0.336	0.115
		C-N	0.1	0.379	(0.363)	0.143	0.948	-16.5	0.330	0.121
		W-N	-12.5	0.340	(0.339)	0.129	0.938	-51.1	0.322	0.224
	W-N	G-N	-0.3	0.341	(0.336)	0.116	0.950	-11.9	0.310	0.102
		C-N	-2.0	0.336	(0.335)	0.113	0.956	-15.5	0.302	0.102
		W-N	0.1	0.340	(0.333)	0.116	0.954	-11.8	0.319	0.108
	W-EV	G-N	2.7	0.355	(0.335)	0.126	0.936	-20.6	0.312	0.116
		C-N	0.5	0.348	(0.334)	0.121	0.944	-24.7	0.304	0.121
		W-N	2.0	0.350	(0.332)	0.123	0.938	-23.4	0.318	0.126
2	G-N	G-N	-0.1	0.438	(0.428)	0.192	0.950	-0.9	0.332	0.110
		C-N	-4.5	0.429	(0.413)	0.185	0.946	-9.7	0.327	0.111
		W-N	-15.6	0.369	(0.370)	0.155	0.916	-42.6	0.318	0.184
	G-EV	G-N	4.0	0.445	(0.430)	0.199	0.940	-7.4	0.343	0.120
		C-N	-1.7	0.426	(0.403)	0.182	0.938	-18.0	0.346	0.134
		W-N	-12.9	0.371	(0.360)	0.152	0.912	-51.3	0.317	0.221
	W-N	G-N	-0.4	0.398	(0.400)	0.158	0.944	-6.9	0.293	0.088
		C-N	-2.7	0.388	(0.394)	0.151	0.944	-12.7	0.295	0.094
		W-N	-1.5	0.389	(0.393)	0.152	0.948	-11.9	0.314	0.105
	W-EV	G-N	1.5	0.396	(0.388)	0.157	0.952	-20.5	0.289	0.103
		C-N	0.4	0.388	(0.382)	0.150	0.950	-26.4	0.288	0.115
		W-N	0.9	0.389	(0.381)	0.151	0.946	-24.6	0.307	0.122

The simulation is conducted with 500 replications using the structure of rat data ($N = 150$ with $n = 50$ & $n_i = 3$, and 73% censoring) in Section 5.1. In design 1, $x_{i1} = 1$ and $x_{i2} = x_{i3} = 0$ for $i = 1, \dots, 50$ and in design 2, $x_{i1} = x_{i2} = x_{i3} = 1$ for the first 25 subjects and 0 for the remaining 25. G-N, C-N, W-N, G-EV and W-EV denote GTDL-normal, Cox's PH-normal, Weibull-normal, GTDL-extreme value and Weibull-extreme value frailty models, respectively. RB% and 95%, respectively, indicate the percentage of relative bias and empirical coverage probability for a nominal 95% confidence interval for β_1 .

Table 5. The true values $(\alpha, \sigma) = (0.15, 1.00)$ and and 73% censoring: simulation results of fitting three frailty models when basic hazard or log-frailty distribution is correctly specified or misspecified. The α and σ are, respectively, a time effect in GTDL model and a square root of log-frailty variance in $N(0, \sigma^2)$. The true regression and Weibull shape parameters are, respectively, $\beta_1 = 0.927$ and $\phi = 1.5$.

Design	True	Fitted	$\hat{\beta}_1$					$\hat{\sigma}$		
			RB%	SD	(SE)	MSE	95%	RB%	SD	MSE
1	G-N	G-N	0.5	0.440	(0.425)	0.193	0.950	-5.5	0.342	0.120
		C-N	-9.4	0.398	(0.382)	0.166	0.948	-12.0	0.336	0.127
		W-N	-18.3	0.367	(0.361)	0.163	0.924	-31.3	0.372	0.236
	G-EV	G-N	4.4	0.457	(0.430)	0.220	0.938	-13.3	0.309	0.113
		C-N	-7.4	0.415	(0.375)	0.176	0.922	-21.8	0.307	0.142
		W-N	-17.3	0.374	(0.354)	0.165	0.924	-39.4	0.339	0.270
	W-N	G-N	-3.7	0.329	(0.344)	0.109	0.956	-13.6	0.270	0.091
		C-N	-4.6	0.329	(0.344)	0.110	0.954	-15.6	0.275	0.100
		W-N	-2.3	0.335	(0.342)	0.113	0.952	-13.3	0.305	0.110
	W-EV	G-N	1.7	0.362	(0.340)	0.131	0.938	-26.2	0.285	0.150
		C-N	-0.2	0.353	(0.339)	0.124	0.944	-29.6	0.282	0.167
		W-N	1.3	0.358	(0.337)	0.128	0.942	-27.4	0.297	0.163
2	G-N	G-N	0.6	0.493	(0.499)	0.243	0.952	-6.3	0.325	0.110
		C-N	-9.3	0.438	(0.453)	0.199	0.940	-13.6	0.322	0.122
		W-N	-17.6	0.396	(0.412)	0.183	0.906	-31.6	0.357	0.227
	G-EV	G-N	3.2	0.495	(0.487)	0.245	0.938	-13.2	0.303	0.109
		C-N	-8.3	0.448	(0.439)	0.206	0.938	-21.2	0.309	0.140
		W-N	-17.4	0.406	(0.399)	0.191	0.902	-39.0	0.338	0.266
	W-N	G-N	-2.0	0.452	(0.437)	0.204	0.946	-14.2	0.265	0.090
		C-N	-3.7	0.444	(0.432)	0.198	0.948	-16.9	0.267	0.100
		W-N	-2.2	0.447	(0.433)	0.200	0.946	-15.1	0.296	0.110
	W-EV	G-N	-2.6	0.413	(0.417)	0.171	0.948	-24.4	0.266	0.130
		C-N	-4.6	0.407	(0.410)	0.167	0.944	-28.5	0.273	0.155
		W-N	-3.0	0.409	(0.412)	0.168	0.954	-25.9	0.282	0.146

The simulation is conducted with 500 replications using the structure of rat data ($N = 150$ with $n = 50$ & $n_i = 3$, and 73% censoring) in Section 5.1. In design 1, $x_{i1} = 1$ and $x_{i2} = x_{i3} = 0$ for $i = 1, \dots, 50$ and in design 2, $x_{i1} = x_{i2} = x_{i3} = 1$ for the first 25 subjects and 0 for the remaining 25. G-N, C-N, W-N, G-EV and W-EV denote GTDL-normal, Cox's PH-normal, Weibull-normal, GTDL-extreme value and Weibull-extreme value frailty models, respectively. RB% and 95%, respectively, indicate the percentage of relative bias and empirical coverage probability for a nominal 95% confidence interval for β_1 .