Statistica Sinica Preprint No: SS-2017-0156	
Title	Regression Analysis of Multivariate Current Status Data with Semiparametric Transformation Frailty Models
Manuscript ID	SS-2017-0156
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202017.0156
Complete List of Authors	Shuwei Li Tao Hu Shishun Zhao and Jianguo Sun
Corresponding Author	Jianguo Sun
E-mail	sunj@missouri.edu
Notice: Accepted version subject to English editing.	

Regression Analysis of Multivariate Current Status Data with Semiparametric Transformation Frailty Models

Shuwei Li[1] Tao Hu[2] Shishun Zhao[3] and Jianguo Sun[4]

[1]School of Economics and Statistics, Guangzhou University, Guangzhou, China

[2]School of Mathematical Sciences, Capital Normal University, Beijing, China

[3]Center for Applied Statistical Research, School of Mathematics, Jilin University, Changchun, China

[4]Department of Statistics, University of Missouri, Columbia, Missouri, USA

Abstract This article investigates regression analysis of multivariate current status data with the use of a class of flexible semiparametric transformation frailty models. For the problem, the maximum likelihood estimation procedure is derived and in particular, a novel EM algorithm, which is quite stable and can be easily implemented, is developed. Also the asymptotic properties of the resulting estimators are established and the numerical study indicates that the proposed methodology works well in practical situations. An application is provided for the illustration of the presented method.

Keywords: EM algorithm; Transformation frailty models; Multivariate current status data; Semiparametric efficiency.

1 Introduction

Current status data, also known as case 1 interval-censored failure time data, occur frequently in many fields such as demographical investigations, epidemiology studies and tumorigenicity experiments (Huang, 1996; Rossini and Tsiatis, 1996; Lin et al., 1998; Martinussen and Scheike, 2002; Jewell and van der Laan, 2004; Xue et al., 2004; Sun, 2006; Zeng et al., 2016). For such data, each subject in the study is observed only once and one only knows that the failure event of interest occurs either before or after the observation time. In other words, the failure time is either left-

Correspondence author: Shishun Zhao, E-mail: zhaoss@jlu.edu.cn

or right-censored and cannot be observed exactly. Multivariate current status data mean that the failure time study involves several correlated failure times of interest and only current status data are available for each of the failure times of interest (Dunson and Dinse, 2002; Jewell et al., 2005; Chen et al., 2009).

A great deal of literature has been developed for regression analysis of univariate current status data and some authors have also discussed regression analysis of multivariate current status data (Sun, 2006). For the latter, however, most of existing methods apply only to some restricted models or limited situations. For the analysis of multivariate failure time data, one of the main challenges is how to deal with the correlation among the correlated failure times. For this, two types of approaches are commonly used in general and they are marginal model-based methods and frailty model-based methods. The former leaves the correlation arbitrary and treats the failure times of interest as independent, which is often referred to as the working independence assumption (Wei et al., 1989; Goggins and Finkelstein, 2000; Chen et al., 2007). The advantage of these methods is their simplicity as, for example, the likelihood function and the corresponding estimation procedure can be relatively simple and easily derived. On the other hand, they may not be efficient compared to the frailty model-based methods (Guo and Rodriguez, 1992).

A frailty model-based approach usually tries to directly model the relationship among the correlated failure times of interest by using the latent variable or frailty and among others, Wen and Chen (2011) and Wang et al. (2015) recently proposed such methods for regression analysis of bivariate current status data under the gamma frailty proportional hazards model. The former developed a nonparametric maximum likelihood technique and the latter employed a spline-based EM algorithm to estimate the parameters involved in the model. Note that the marginal approach aims to estimate the population-average covariate effect, while the frailty approach allows one to estimate the subject-specific effects. Also note that it is well-known that the proportional hazards model may not provide proper fit sometimes. In the following, we will develop a frailty model-based approach with the use of a class of flexible semiparametric transformation frailty models. In addition to the differences discussed above between univariate and multivariate current status data, it is clear that the multivariate data also have much more complex structures.

The remainder of the paper is organized as follows. First in Section 2, we introduce some notation and assumptions that will be used throughout the paper. The semiparametric transformation frailty models are then described along with the resulting likelihood function. Section 3 provides the nonparametric maximum likelihood estimation procedure and for the implementation of the

procedure, a novel EM algorithm is developed with the use of some Poisson latent variables. In particular, the algorithm employs the probability integral transformation technique and Gauss-Hermite quadrature method together in the E-step. In Section 4, the asymptotic properties of the resulting estimators, including the consistency, asymptotical normality and semiparametric efficiency, are established and Section 5 presents some results obtained from a simulation study, which suggest that the proposed methodology works well for practical situations. In Section 6, we illustrate the proposed method with a real data example, and Section 7 contains some discussion and concluding remarks.

2 Notation, Assumptions and the Likelihood Function

Consider a failure time study that involves n independent subjects and each subject can experience K possibly correlated failure events of interest. For subject i, let T_{ik} denote the failure time of the kth event and X_{ik} be the corresponding d-dimensional vector of covariates, $i = 1, ..., n$. Suppose that for each T_{ik}, only one observation is available at the observation time C_{ik} and we only know that the event occurs either before or after C_{ik}. In other words, T_{ik} is either left- or right-censored at C_{ik} and the observed data have the form $O_{ik} = \{ C_{ik}, \Delta_{ik} = I(T_{ik} \leq C_{ik}), X_{ik} \}$ with $I(\cdot)$ being the indicator function. In the following, we will assume that T_{ik} and C_{ik} are conditionally independent given the covariate X_{ik}.

To describe the covariate effects, we will assume that there exists a latent variable b_i and given X_{ik} and b_i, the cumulative hazard function of T_{ik} has the form

$$G_k \left\{ \Lambda_k(t)\, e^{X_{ik}^T \beta}\, b_i \right\}, \tag{1}$$

where $\Lambda_k(t)$ denotes an unknown baseline cumulative hazard function, β is a d-dimensional vector of regression parameters, and G_k is a prespecified increasing function. Note that many authors, including Dabrowska and Doksum (1988) and Zeng and Lin (2007), have discussed the same or similar models and it is easy to see that this class of models contains many commonly used models as special cases. For example, by letting $G_k(x) = x$, we obtain the proportional hazards frailty model and it gives the proportional odds frailty model with $G_k(x) = \log(1 + x)$. Also note that in the class of models (1), we have assumed that the covariate effects are same for different failure times for simplicity of presentation. In the case that they may be different, one can still apply the methodology proposed below by simply defining a new, larger vector of covariates. In the

following, we will assume that given b_i, $T_{i1}, \ldots, T_{iK}$ are independent of each other and the b_i's follow a parametric model with mean one and the density function $p(b_i \mid \gamma)$, where γ is an unknown parameter. Then the likelihood function has the form

$$L(\beta, \gamma, \Lambda) = \prod_{i=1}^{n} \int_{b_i} \prod_{k=1}^{K} \left\{1 - \exp\left[-G_k\left\{\Lambda_k(C_{ik})\, e^{X_{ik}^T \beta}\, b_i\right\}\right]\right\}^{\Delta_{ik}}$$

$$\times \exp\left[-G_k\left\{\Lambda_k(C_{ik})\, e^{X_{ik}^T \beta}\, b_i\right\}\right]^{1-\Delta_{ik}} p(b_i \mid \gamma)\, \mathrm{d}b_i$$

with $\Lambda = (\Lambda_1, ..., \Lambda_K)$. To maximize $L(\beta, \gamma, \Lambda)$, it is apparent that one has to perform some numerical integration and the direct maximization would be quite challenging and unstable even under cox model setting (Wang et al., 2015). More importantly, the resulting estimators have no closed forms, which naturally suggests the use of the following EM algorithm.

Note that as discussed by Kosorok et al. (2004), in model (1), the transformation function $G_k(x)$ can be derived by or written in the following Laplace transformation form

$$\exp\{-G_k(x)\} = \int_0^\infty e^{-xt}\, \phi(t \mid r_k)\, \mathrm{d}t\,,$$

where $\phi(t \mid r_k)$ is a density function depending some constant r_k with the support $[0, \infty)$. An example of $\phi(t \mid r_k)$ is the gamma density function with mean one and variance r_k, which yields $G_k(x) = \log(1 + r_k\, x)/r_k$, the logarithmic transformation function. One advantage of the latter form is that one can convert the transformation frailty model into the proportional hazards model with two sets of random effects. More specifically, let μ_{ik} denote the latent variable following the density function $\phi(t \mid r_k)$. Then the conditional survival function of T_{ik} can be expressed as

$$S_k(t \mid X_{ik}, b_i) = \int_{\mu_{ik}} \exp\left[-\mu_{ik}\left\{\Lambda_k(t)\, e^{X_{ik}^T \beta}\, b_i\right\}\right]\, \phi(\mu_{ik} \mid r_k)\, \mathrm{d}\mu_{ik}$$

given X_{ik} and b_i. It follows that the likelihood function $L(\beta, \gamma, \Lambda)$ can be rewritten as

$$L_1(\beta, \gamma, \Lambda) = \prod_{i=1}^{n} \int_{b_i} \prod_{k=1}^{K} \int_{\mu_{ik}} \left\{1 - \exp\left[-\mu_{ik}\left\{\Lambda_k(C_{ik})\, e^{X_{ik}^T \beta}\, b_i\right\}\right]\right\}^{\Delta_{ik}}$$

$$\times \exp\left[-\mu_{ik}\left\{\Lambda_k(C_{ik})\, e^{X_{ik}^T \beta}\, b_i\right\}\right]^{1-\Delta_{ik}} \phi(\mu_{ik} \mid r_k)\, \mathrm{d}\mu_{ik}\, p(b_i \mid \gamma)\, \mathrm{d}b_i\,. \tag{2}$$

In the next section, we will discuss the estimation of (β, γ, Λ) based on $L_1(\beta, \gamma, \Lambda)$ given in (2).

3 Maximum Likelihood Estimation

Now we discuss the estimation of (β, γ, Λ) and for this, we will derive the nonparametric maximum likelihood estimation procedure. For each k, let $t_{1k} < \ldots < t_{n_k k}$ denote the distinct ordered observation times of $\{ C_{ik}; i = 1, \ldots, n \}$ and assume that Λ_k is a step function with nonnegative jump size λ_{lk} at t_{lk}, $l = 1, \ldots, n_k$. In other words, we have $\Lambda_k(t) = \sum_{t_{lk} \le t} \lambda_{lk}$. Let θ be all unknown parameters to be estimated, then we can rewrite the likelihood function (2) as the following form

$$L_2(\theta) = \prod_{i=1}^{n} \int_{b_i} \prod_{k=1}^{K} \int_{\mu_{ik}} \left\{ 1 - \exp\left[-\mu_{ik} \left(\sum_{t_{lk} \le C_{ik}} \lambda_{lk} \right) e^{X_{ik}^T \beta}\, b_i \right] \right\}^{\Delta_{ik}}$$

$$\times \exp\left[-\mu_{ik} \left(\sum_{t_{lk} \le C_{ik}} \lambda_{lk} \right) e^{X_{ik}^T \beta}\, b_i \right]^{1-\Delta_{ik}} \phi(\mu_{ik} \mid r_k)\, \mathrm{d}\mu_{ik}\, p(b_i \mid \gamma)\, \mathrm{d}b_i.$$

In the following, we will develop an EM algorithm based on a two-stage data augmentation with the use of Poisson variables.

In the first stage, we will assume that the latent variables b_i's and μ_{ik}'s were known, and in this case, the likelihood function would have the form

$$L_3(\theta) = \prod_{i=1}^{n} \prod_{k=1}^{K} \left\{ 1 - \exp\left[-\mu_{ik} \left(\sum_{t_{lk} \le C_{ik}} \lambda_{lk} \right) e^{X_{ik}^T \beta}\, b_i \right] \right\}^{\Delta_{ik}}$$

$$\times \exp\left[-\mu_{ik} \left(\sum_{t_{lk} \le C_{ik}} \lambda_{lk} \right) e^{X_{ik}^T \beta}\, b_i \right]^{1-\Delta_{ik}} \phi(\mu_{ik} \mid r_k)\, p(b_i \mid \gamma)\,.$$

In the second stage, define a mapping between Δ_{ik} and a new latent variable Z_{ik} by $\Delta_{ik} = I(Z_{ik} > 0)$, where $Z_{ik} = \sum_{t_{lk} \le C_{ik}} Z_{ilk}$ with Z_{ilk}'s being the independent Poisson random variables with means $\mu_{ik} \lambda_{lk} e^{X_{ik}^T \beta}\, b_i$ $(i = 1, \ldots, n;\ k = 1, \ldots, K;\ l = 1, \ldots, n_k)$. Hence if the Z_{ilk}'s were known, we would have the following complete data likelihood function

$$L_c(\theta) = \prod_{i=1}^{n} \prod_{k=1}^{K} \prod_{l=1}^{n_k} \psi(Z_{ilk} \mid \mu_{ik} \lambda_{lk} e^{X_{ik}^T \beta}\, b_i)\, \phi(\mu_{ik} \mid r_k)\, p(b_i \mid \gamma)$$

subject to the constrains that $Z_{ik} = \sum_{t_{lk} \le C_{ik}} Z_{ilk} > 0$ if $\Delta_{ik} = 1$ and $Z_{ik} = \sum_{t_{lk} \le C_{ik}} Z_{ilk} = 0$ if $\Delta_{ik} = 0$. Here, $\psi(Z_{ilk} \mid \mu_{ik} \lambda_{lk} e^{X_{ik}^T \beta}\, b_i)$ is the probability mass function of Z_{ilk} with the parameter being $\mu_{ik} \lambda_{lk} e^{X_{ik}^T \beta}\, b_i$. Of course, by integrating out the latent variables Z_{ilk}'s, $L_c(\theta)$ will reduce back to $L_3(\theta)$.

Let $\theta^{(m)}$ denote the estimator of θ obtained in the mth iteration. To obtain $\theta^{(m+1)}$, in the E-step, we need to take the logarithm of complete data likelihood function $L_c(\theta)$ and calculate the following conditional expectations with respect to all latent variables

$$\mathrm{E}(\mu_{ik}\, b_i) = \mathrm{E}_{b_i}\left\{ b_i \, \frac{\Delta_{ik} - \exp(-G_k(W_{ik}))G_k'(W_{ik})}{\Delta_{ik} - \exp(-G_k(W_{ik}))} \right\},$$

$$\mathrm{E}(Z_{ilk}) = \Delta_{ik}\, \lambda_{lk}\, e^{X_{ik}^T\beta}\, \mathrm{E}_{b_i}\left\{ \frac{b_i}{1-\exp(-G_k(W_{ik}))} \right\} I(t_{lk} \le C_{ik}) + \lambda_{lk}\, e^{X_{ik}^T\beta}\, \mathrm{E}(\mu_{ik}\, b_i\,)\, I(t_{lk} > C_{ik}),$$

and

$$\mathrm{E}\{h(b_i)\} = \frac{\int_{b_i} h(b_i) \prod_{k=1}^K \{1-\exp(-G_k(W_{ik}))\}^{\Delta_{ik}} \exp\{-G_k(W_{ik})\}^{1-\Delta_{ik}}\, p(b_i \mid \gamma)\, \mathrm{d}b_i}{\int_{b_i} \prod_{k=1}^K \{1-\exp(-G_k(W_{ik}))\}^{\Delta_{ik}} \exp\{-G_k(W_{ik})\}^{1-\Delta_{ik}}\, p(b_i \mid \gamma)\, \mathrm{d}b_i}.$$

In the above, $h(b_i)$ is an arbitrary function of b_i, $W_{ik} = \sum_{t_{lk} \le C_{ik}} \lambda_{lk}\, e^{X_{ik}^T\beta}\, b_i$ and

$$G'(W_{ik}) = \frac{\int_{\mu_{ik}} \mu_{ik} e^{-W_{ik}\,\mu_{ik}}\, \phi(\mu_{ik} \mid r_k)\, \mathrm{d}\mu_{ik}}{\exp\{-G_k(W_{ik})\}}.$$

For notational simplicity, in the above, we have suppressed the conditional arguments in all conditional expectations. Also note that if $\phi(\mu_{ik} \mid r_k)$ is the gamma density function, the integration above with respect to μ_{ik} has the following closed form

$$\int_{\mu_{ik}} \mu_{ik} e^{-W_{ik}\,\mu_{ik}}\, \phi(\mu_{ik} \mid r_k)\, \mathrm{d}\mu_{ik} = (r_k\, W_{ik} + 1)^{-r_k^{-1}-1}.$$

Otherwise we suggest to employ the Gaussian-Laguerre quadrature technique to calculate the integration with respect to μ_{ik}. For determining $\mathrm{E}\{h(b_i)\}$, we suggest to employ the probability integral transformation technique to transform b_i into a standard normal random variable and then adopt Gaussian-Hermite quadrature method. The numerical study below suggests that the joint use of the probability integral transformation and Gaussian-Hermite quadrature techniques performs well in practice. Nelson et al. (2006) gives a detailed discussion about the probability integral transformation when the random effects or frailties follow non-normal distributions.

In the M-step, we need to maximize the following objective function with respect to θ

$$Q(\theta, \theta^{(m)}) = \sum_{i=1}^n \sum_{k=1}^K \sum_{l=1}^{n_k} \left\{ X_{ik}^T\beta\, \mathrm{E}(Z_{ilk}) + \log(\lambda_{lk})\mathrm{E}(Z_{ilk}) - \lambda_{lk} e^{X_{ik}^T\beta}\, \mathrm{E}(\mu_{ik}\, b_i) \right\}$$
$$+ \sum_{i=1}^n \mathrm{E}\{\log(p(b_i \mid \gamma))\}.$$

Setting $\partial Q(\theta, \theta^{(m)}) \,/\, \partial \lambda_{lk} = 0$, we can update λ_{lk} with the following closed-form expression

$$\lambda_{lk} = \frac{\sum_{i=1}^n \mathrm{E}(Z_{ilk})}{\sum_{i=1}^n \mathrm{E}(\mu_{ik}\, b_i)\, e^{X_{ik}^T\beta}},\ k = 1, \ldots, K,\ l = 1, \ldots, n_k. \tag{3}$$

By plugging the estimators above into $Q(\theta, \theta^{(m)})$, we can obtain the score equations for β as

$$\sum_{i=1}^{n}\sum_{k=1}^{K}\left\{\left(\sum_{l=1}^{n_k}\mathrm{E}(Z_{ilk})\right)\left(X_{ik}-\frac{\sum_{i=1}^{n}\mathrm{E}(\mu_{ik}\,b_i)e^{X_{ik}^T\beta}X_{ik}}{\sum_{i=1}^{n}\mathrm{E}(\mu_{ik}\,b_i)e^{X_{ik}^T\beta}}\right)\right\}=0\,. \tag{4}$$

Finally, by setting $\partial Q(\theta, \theta^{(m)})\,/\,\partial\gamma = 0$, the estimator of γ can be obtained by solving the score equation

$$\sum_{i=1}^{n}\partial\mathrm{E}\{\log(p(b_i\mid\gamma))\}\,/\,\partial\gamma=0\,.$$

In summary, by combining all steps above, the EM algorithm can be described as follows.

Step 0. Choose an initial estimate $\theta^{(0)}$.

Step 1. At the $(m+1)$th iteration, first calculate the conditional expectations $\mathrm{E}(\mu_{ik}\,b_i)$, $\mathrm{E}(Z_{ilk})$ and $\mathrm{E}\{h(b_i)\}$ at $\theta^{(m)}$.

Step 2. Update $\beta^{(m+1)}$ by solving the equation (4) with the use of one-step Newton-Raphson method.

Step 3. Obtain $\lambda_{lk}^{(m+1)}$ by the expression (3).

Step 4. Calculate $\gamma^{(m+1)}$ by solving $\sum_{i=1}^{n}\partial\mathrm{E}\{\log(p(b_i\mid\gamma))\}\,/\,\partial\gamma=0$.

Step 5. Repeat Steps 1 - 4 until the convergence is achieved.

In the above estimation procedure, we have assumed that the constants r_k's are known as they are generally unidentifiable without other assumptions or extra data (Zeng and Lin, 2007). In practice, one common way for their determination is to try different values for them and then select the best or optimal one based on some criterion such as the maximum likelihood principle. More comments on this are given below.

4 Asymptotic Properties

Let $\zeta=(\beta^T,\gamma)^T$, and $\zeta_0=(\beta_0^T,\gamma_0)^T$ and $\theta_0=(\zeta_0^T,\Lambda_{10},...,\Lambda_{K0})$ denote the true values of ζ and θ, respectively. Also let $\hat\theta_n=(\hat\zeta_n^T,\hat\Lambda_{1n},...,\hat\Lambda_{Kn})$ denote the maximum likelihood estimator of θ defined in the previous section. In the following, we will establish the asymptotic properties of $\hat\theta_n$ and for this, we will first present some needed regularity conditions.

(A1) The true value ζ_0 belongs to a known compact set $\mathcal{A}\otimes\mathcal{B}$ in R^{d+1}. Also given covariates, each examination time C_k has a continuous conditional density function with the support $[\tau_1,\tau_2]$ and the true value $\Lambda_{k0}(\cdot)$ is continuously differentiable with positive derivatives in $[\tau_1,\tau_2]$ with $M^{-1}<\Lambda_{k0}(\tau_1)<\Lambda_{k0}(\tau_2)<M$, for $k=1,\ldots,K$, where M is a large positive constant.

(A2) The covariate vectors X_k's are bounded.

(A3) The transformation function G_k is twice continuously differentiable on $[0, \infty)$ with $G_k(0) = 0$, $G_k'(x) > 0$ and $G_k(\infty) = \infty$.

(A4) For any smooth function $g(\cdot)$, we have $\sup_{\gamma \in \mathcal{C}} \int_b g(b)\, p^{(j)}(b \mid \gamma)\, \mathrm{d}b < \infty$ for $j = 0, 1, 2$, where $p^{(j)}(b \mid \gamma)$ denotes the jth derivative of $p(b \mid \gamma)$ with respect to γ.

(A5) There exist $c_1, \ldots, c_K \in [\tau_1, \tau_2]$ for which there are $d + K + 1$ different values of $(\Delta_1, \ldots, \Delta_K, X_1, \ldots, X_K)$ such that if

$$\left(u^T \frac{\partial}{\partial \zeta} + \sum_{k=1}^{K} v_k \frac{\partial}{\partial y_k} \right) \Bigg|_{(\zeta, y_1, \ldots, y_K) = (\zeta_0, \Lambda_{10}(c_1), \ldots, \Lambda_{K0}(c_K))}$$

$$\log \int_b \prod_{k=1}^{K} \left\{ \Delta_k + (-1)^{\Delta_k} \exp\left[-G_k\left(y_k\, e^{X_k^T \beta}\, b \right) \right] \right\} p(b \mid \gamma)\, \mathrm{d}b = 0$$

for each of these $d + K + 1$ values, then $u = 0_{(d+1)\times 1}$ and $v_k = 0$. Here $0_{(d+1)\times 1}$ denotes a $(d+1)$-dimensional vector of zeros.

The conditions above are mild and can be satisfied in practical situations. Conditions (A1) and (A2) are standard conditions in survival analysis. Condition (A3) pertains to the transformation function and it is easy to check that it holds for the logarithmic transformation function $G_r(x) = r^{-1} \log(1 + rx)$ $(r \geq 0)$ among others. Also condition (A4) is commonly required for modeling the multivariate data with frailty models, and condition (A5) is needed for the identifiability of the model (Chang et al. 2007). Now we are ready to present the asymptotic properties of $\hat{\theta}_n$. In the following, let $\| \cdot \|$ be the Euclidean norm, and for a function f and a random variable X with the distribution P, define $\mathbb{P}f = \int f(x) \mathrm{d}P(x)$ and $\mathbb{P}_n f = n^{-1} \sum_{i=1}^{n} f(X_i)$.

Theorem 1 (Consistency). Suppose that conditions (A1)-(A5) hold, we have that as $n \to \infty$, $\|\hat{\zeta}_n - \zeta_0\| \to 0$ and $\sum_{k=1}^{K} \sup_{t \in [\tau_1, \tau_2]} |\hat{\Lambda}_{kn}(t) - \Lambda_{k0}(t)| \to 0$ in probability.

Theorem 2 (Rate of convergence). Suppose that conditions (A1)-(A5) hold, we have that as $n \to \infty$, $d(\hat{\theta}_n, \theta_0) = \left\{ \|\hat{\zeta}_n - \zeta_0\|^2 + \sum_{k=1}^{K} \int [\hat{\Lambda}_{kn}(c) - \Lambda_{k0}(c)]^2 f_k(c) \mathrm{d}c \right\}^{1/2} = O_p(n^{-1/3})$, where $f_k(c)$ denotes the density of C_k.

Theorem 3 (Asymptotic normality and efficiency). Suppose that conditions (A1)-(A5) hold, we have that as $n \to \infty$, $\sqrt{n}(\hat{\zeta}_n - \zeta_0) \xrightarrow{d} N(0, I_0^{-1})$, where $I_0 = \mathbb{P}\{\tilde{l}(\theta_0)\tilde{l}(\theta_0)^T\}$ with $\tilde{l}(\theta_0)$, given in the online supplementary document, denoting the efficient score for ζ at θ_0.

The theorems above say that the maximum likelihood estimator $\hat{\zeta}_n$ is asymptotically efficient and the estimators $\hat{\Lambda}_{kn}$'s only have the $n^{-1/3}$ convergence rate. The proofs of the results above are sketched in the online supplementary document. To make inference about β and γ based on the theorems above, it is apparent that we need to estimate the asymptotic covariance matrix of the corresponding estimators. Since it would be very difficult to derive the consistent estimator of I_0^{-1}, we suggest to employ the nonparametric bootstrap method (Efron, 1981; Su and Wang, 2016) as follows. Let Q be an integer and for $1 \le q \le Q$, draw a new data set, denoted by of $Q^{(q)}$, of sample size n with replacement from the original observed data $O = (O_i = (O_{i1}, ..., O_{iK}); i = 1, \ldots, n)$. Let $\hat{\beta}_n^{(q)}$ and $\hat{\gamma}_n^{(q)}$ denote the maximum likelihood estimators of β_0 and γ_0 defined above based on the bootstrap sample $Q^{(q)}$, respectively. Then one can estimate the covariance matrix and variance of $\hat{\beta}_n$ and $\hat{\gamma}_n$ by using the sample covariance matrix and variance of the $\hat{\beta}_n^{(q)}$'s and $\hat{\gamma}_n^{(q)}$'s, respectively. The numerical studies below indicate that this method works well for practical situations.

5 A Simulation Study

In this section, we report some results obtained from an extensive simulation study performed to investigate the finite sample performance of the methodology proposed in the previous sections. In the study, we considered the situation where there exist $K = 2$ correlated failure times and for simplicity, we assumed that $C_{i1} = C_{i2}$ with the observation times generated from the uniform distribution over $(3, 5)$. Also it was assumed that $X_{i1} = X_{i2}$ and there exist two covariates with the first covariate generated from the Bernoulli distribution with the success probability of 0.5 and the second covariate following the uniform distribution over $(0, 1)$. To generate the failure times, we took G_k to be the logarithmic transformation function and $\Lambda_k(t) = 0.05\, t^2$, and supposed that the latent variables b_i's follow the log-normal distribution with mean one and variance γ_0^2 or the gamma distribution with mean one and variance γ_0. The results given below are based on 1000 replications with $Q = 50$ and $n = 200$ or 400.

Table 1 presents the results obtained on estimation of β and γ with $(\beta_{10}, \beta_{20}) = (0, 0.5)$ or $(0.5, -0.5)$ and $\gamma_0 = 1$. They include the estimated bias (Bias) given by the average of the estimates minus the true value, the sample standard error (SSE) of the estimates, the average of the standard error estimates (SEE), and the 95% empirical coverage probability (CP). It can be seen from the table that the proposed maximum likelihood estimators seem to be unbiased and the bootstrap variance estimates are appropriate. In addition, the normal approximation to

the distribution of the estimators appears to be reasonable and as expected, the results become better when the sample size increases. Furthermore, the estimation procedure seems to give similar performance for different G_k. We also considered some other set-ups, including the situations with different assumed functions for $\Lambda_1(t)$ and $\Lambda_2(t)$, the other types of G_k, or different distribution functions for the b_i's, and obtained similar results.

Note that in the proposed methodology, it has been assumed that the distribution for the b_i's is known and thus one question of interest is the robustness of the estimation procedure to the misspecification of the latent variable distribution. To assess this, we repeated the study above in which we generated the b_i's from the gamma distribution with mean one and variance one but assumed that they were from the log-normal distribution. Table 2 gives the estimation results on β with $n = 200$ and the other model specification being the same as that in Table 1. One can see that as with Table 1, the proposed estimators again seem to perform well and the results suggest that the estimation procedure appears to be robust with respect to the latent variable distribution. For the question here, we also studied the situations such that the b_i's were generated from the log-normal distribution but wrongly assumed to be from the gamma distribution and obtained similar results.

In the simulation study, we also compared the proposed method to that given in Wang et al. (2015), who discussed regression analysis of bivariate current status data with a special case of transformation models $G_k\left\{\Lambda_k(t)\, e^{X_{ik}^T\beta^{(k)}}\, b_i\right\}$ with $G_1(x) = G_2(x) = x$ and the b_i's following the gamma distribution. For comparison, we repeated the study giving results in Table 1 but with the b_i's generated from the gamma distribution with both mean and variance being one and one common covariate following the Bernoulli distribution, and presented the obtained results in Table 3. Note that here we only considered the estimated bias and the sample standard error on the estimation of $\beta = (\beta^{(1)}, \beta^{(2)})^T$ with $n = 200$, various combination of true values for regression parameters $(\beta_0^{(1)}, \beta_0^{(2)})^T$, and $\gamma_0 = 0.5$, 1 or 1.5. Also here we assumed that the covariate effect may be different for two different failure times and as mentioned above, the proposed method can directly apply to this situation. It is apparent that the two methods give similar results and in particular, they have similar efficiency.

6 An Illustration

Now we apply the proposed methodology to a set of real bivariate current status data arising from the Infertility Prevention Project, which was designed through the screening test of the risk subjects to asses the prevalence of chlamydia and gonorrhea throughout the United States. The chlamydia and gonorrhea are two sexually transmitted diseases that can frequently co-exist and lead to complicated clinical syndromes if left untreated. The data set considered here consists of 5879 subjects in Nebraska whose urine specimens were collected during the individuals' visits to health clinics in 2008 and then sent to the Nebraska Public Health Laboratory (NPHL) for testing the infection status of both diseases. For the data, the overall prevalence of chlamydia and gonorrhea is approximately 0.083 and 0.017, respectively, and the factors or covariates of interest include the gender, whether the patient is Caucasian, and whether the patient has some symptoms at the test. For the analysis, as in many epidemiological surveys, we will focus on the ages of the subjects at the infection of the two diseases with the age at test serving as the observation time. In other words, we have $C_{i1} = C_{i2}$ for the data here.

To apply the proposed method, let T_1 denote the age of chlamydia infection and T_2 be the age of gonorrhea infection. Also let Gender (1 for male and 0 for female), C-America (1 for yes and 0 for no) and Symptoms (1 for yes and 0 for no) represent the three covariates described above. As in the simulation study, we employed both log-normal and gamma frailty distributions for the latent variables b_i's and the logarithmic transformation function. We considered the equally spaced grid points of r_1 and r_2 ranging from 0 to 3 with the increments of 0.1 for the functions G_k's and then the maximum likelihood principle was used for the selection of the optimal model. Through the analysis, under the log-normal and gamma frailty distributions, the optimal model was given by $(r_1, r_2) = (2.1, 2.6)$ and $(r_1, r_2) = (1.3, 1.5)$, respectively. Table 4 gives the estimated covariate effects obtained under the optimal model, and for comparison, we also included in the table the corresponding results obtained under the proportional hazards (PH) frailty model and the proportional odds (PO) frailty model. The results include the point estimates, the estimated standard errors and the p-values for testing no covariate effect for each of the three covariates.

First one can see from Table 4 that the results are quite consistent across the three models and the two frailty distributions. They all suggest that it seems that the Caucasians have significant lower risks of developing chlamydia and gonorrhea infections than other races, and the patients with the symptoms are more likely to develop the infections than those without the symptoms. However, the risk of developing chlamydia and gonorrhea infections does not seem to be significantly related to the gender of the patients. In addition, under the optimal model with the log-normal frailty

distribution, we got $\hat{\gamma}_n = 0.518$ and this indicates that there may exist some correlation between the two failure events. The similar results were obtained under the other models and the frailty distribution.

Note that as mentioned above, with model (1), we have assumed that the three covariates had the same effects on the two events. Suggested by a reviewer, we also performed the analysis by assuming that the effects may be different and presented the obtained results in Table 5. It is apparent that the overall results and conclusions are similar to those given above. On the other hand, it is clear that the covariates Caucasian and Symptoms had much stronger effects on the risk of developing the gonorrhea infection than on the risk of developing the chlamydia infection. To further see this, we obtained and presented in Figure 1 the estimates of the baseline cumulative hazard functions for chlamydia and gonorrhea infections under the optimal model with the log-normal frailty. Again it indicates that the two baseline hazards seem to be quite different. Note that in the above analysis, we set bootstrap sample size $Q = 50$ as in the simulation study, and we also considered other values for Q and obtained similar results.

7 Discussion and Concluding Remarks

This paper discussed regression analysis of multivariate current status data under a class of flexible semiparametric transformation frailty models and as mentioned above, it includes many existing, specific models as special cases. For inference, the nonparametric maximum likelihood procedure was derived and for the implementation of the procedure, a novel EM algorithm was developed with the use of some Poisson random variables and can be easily implemented. Also the asymptotic properties of the resulting estimators were established and in particular, the estimators of regression parameters were shown to be efficient. In addition, numerical studies were conducted and suggested that the proposed methodology works well in practical situations.

It is worth pointing out that one of the distinct features for the developed EM algorithm is the joint use of the probability integral transformation technique and Gauss-Hermite quadrature method, which allow one to easily calculate the conditional expectations involving various frailty distributions. The use of some Poisson variables allows the high-dimensional parameters λ_{lk}'s to be calculated explicitly and the low-dimensional parameters β and γ to be updated with one-step Newton-Raphson method separately. In consequence, this avoids the inversion of the possibly high-dimensional matrix and hence makes the estimating procedure computationally stable.

The focus of the discussion above has been on the situation of time-independent covariates and it is apparent that sometimes there may exist time-dependent covariates. It is straightforward to generalize the idea and method discussed above to this latter situation although one will need to reformulate the class of models (1) in a different way. Also in the above sections, we have assumed that the r_k's are known and it would be helpful to develop some simultaneous estimation procedures. However, as mentioned above, this is usually not possible without some assumptions or extra information and one can employ some selection criterion for their determination in practice. It should be noted that according to simulation studies, the misspecification of them could cause some mild biases, especially on estimation of γ.

Acknowledgements.

The authors wish to thank the Co-Editor, Dr. Zhiliang Ying, an Associate Editor and two reviewers for their many helpful comments that greatly improved the paper. This work was partly supported by the National Nature Science Foundation of China grant nos. 11671274, 11731011 and 11671168, the Support Project of High-level Teachers in Beijing Municipal Universities in the Period of 13th Five-year Plan grant CIT &TCD 201804078, the Capacity Building for Sci-Tech Innovation - Fundamental Scientific Research Funds grant 025185305000/204, the Youth Innovative Research Team of Capital Normal University, and the Science and Technology Developing Plan of Jilin Province grant 20170101061JC.

References

[1] Chang, I. -S., Wen, C. -C. and Wu, Y. -J. (2007). A profile likelihood theory for the correlated gamma-frailty model with current status family data. *Statistica Sinica*, **17**, 1023-1046.

[2] Chen, M. H., Tong, X. W. and Sun, J. (2007). The proportional odds model for multivariate interval-censored failure time data. *Statistics in Medicine*, **26**, 5147-5161.

[3] Chen, M. H., Tong, X. W. and Sun, J. (2009). A frailty model approach for regression analysis of multivariate current status data. *Statistics in Medicine*, **28**, 3424-3436.

[4] Dabrowska, D. M. and Doksum, K. A. (1988). Partial likelihood in transformation models with censored data. *Scandinavian Journal of Statistics*, **15**, 1-23.

[5] Dunson, D. B. and Dinse, G. E. (2002). Bayesian models for multivariate current status data with informative censoring. *Biometrics*, **58**, 79-88.

[6] Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, **76**, 312-319.

[7] Guo, G. and Rodriguez, G. (1992). Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala. *Journal of the American Statistical Association*, **87**, 969-976.

[8] Goggins, W. B. and Finkelstein, D. M. (2000). A proportional hazards model for multivariate interval-censored failure time data. *Biometrics*, **56**, 940-943.

[9] Huang, J. (1996). Efficient estimation for the Cox model with interval censoring. *The Annals of Statistics*, **24**, 540-568.

[10] Jewell, N. P. and van der Laan, M. J. (2004). Current status data: review, recent developments and open problems. *Advances in Survival Analysis*, Ellsevier, Amsterdam, 625-642.

[11] Jewell, N. P., van der Laan, M. J. and Lei, X. (2005). Bivariate current status data with univariate monitoring times. *Biometrika*, **92**, 847-862.

[12] Kosorok, M. R., Lee, B. L. and Fine, J. P. (2004). Robust inference for univariate proportional hazards frailty regression models. *The Annals of Statistics*, **32**, 1448-1491.

[13] Lin, D. Y., Oakes, D. and Ying, Z. (1998). Additive hazards regression with current status data. *Biometrika*, **85**, 289-298.

[14] Martinussen, J. and Scheike, T. H. (2002). Efficient estimation in additive hazards regression with current status data. *Biometrika*, **89**, 649-658.

[15] Nelson, K. P., Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J., Parzen, M., and Strawderman, R.(2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics* **15**, 39-57.

[16] Rossini, A. J. and Tsiatis, A. A. (1996). A semiparametric proportional odds regression model for the analysis of current status data. *Journal of the American Statistical Association*, **91**, 713-721.

[17] Su, Y.-R. and Wang, J. -L. (2016). Semiparametric efficient estimation for shared-frailty models with doubly-censored clustered data. *The Annals of Statistics* **44**, 1298-1331.

[18] Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data*. Springer, New York.

[19] van der Vaart, A. W., and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

[20] van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Univ. Press, New York.

[21] Wang, N., Wang, L. and McMahan, C. S. (2015). Regression analysis of bivariate current status data under the Gamma-frailty proportional hazards model using the EM algorithm. *Computational Statistics and Data Analysis*, **83**, 140-150.

[22] Wei, L. J., Lin, D. Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, **84**, 1065-1073.

[23] Wen, C. -C. and Chen, Y. -H. (2011). Nonparametric maximum likelihood analysis of clustered current status data with the gamma-frailty Cox model. *Computational Statistics and Data Analysis*, **55**, 1053-1060.

[24] Xue, H., Lam, K. F. and Li, G. (2004). Sieve maximum likelihood estimator for semiparametric regression models with current status data. *Journal of the American Statistical Association*, **99**, 346-356.

[25] Zeng, D. and Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society: Series B*, **69**, 507-564.

[26] Zeng, D., Mao, L., and Lin, D. Y. (2016). Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika*, **103**, 253-271.

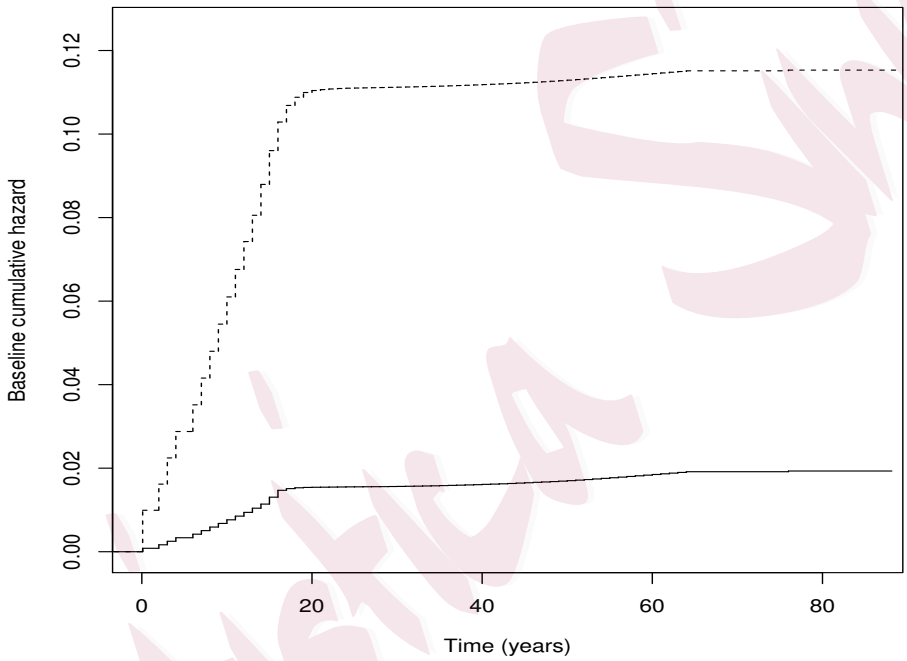

Fig. 1 The estimated baseline cumulative hazard functions for the chlamydia infection (upper step function) versus the gonorrhea infection.

Table 1. Estimation of regression and variance parameters with the log-normal latent variable distribution

n	(β_{10}, β_{20})	Par	Bias	SSE	SEE	CP	(β_{10}, β_{20})	Par	Bias	SSE	SEE	CP
					$G_1(x) = x$ & $G_2(x) = x$							
200	$(0, 0.5)$	β_1	-0.011	0.197	0.203	97.8	$(0.5, -0.5)$	β_1	-0.014	0.224	0.216	94.8
		β_2	-0.012	0.346	0.346	95.2		β_2	-0.016	0.366	0.363	94.9
		γ	0.022	0.174	0.170	94.9		γ	0.021	0.186	0.173	94.0
400	$(0, 0.5)$	β_1	0.006	0.150	0.150	95.0	$(0.5, -0.5)$	β_1	0.007	0.157	0.168	96.0
		β_2	-0.010	0.284	0.265	94.6		β_2	-0.012	0.286	0.285	95.0
		γ	0.012	0.136	0.130	95.0		γ	0.005	0.133	0.132	95.0
					$G_1(x) = 2\log(1 + x/2)$ & $G_2(x) = 2\log(1 + x/2)$							
200	$(0, 0.5)$	β_1	-0.005	0.245	0.250	95.8	$(0.5, -0.5)$	β_1	0.009	0.273	0.269	95.0
		β_2	0.021	0.441	0.446	95.1		β_2	0.024	0.419	0.432	96.8
		γ	0.030	0.245	0.260	95.2		γ	0.025	0.244	0.248	95.2
400	$(0, 0.5)$	β_1	-0.004	0.169	0.172	95.2	$(0.5, -0.5)$	β_1	0.006	0.185	0.182	94.9
		β_2	0.009	0.310	0.311	95.2		β_2	0.014	0.305	0.300	95.2
		γ	0.023	0.177	0.177	95.0		γ	0.017	0.170	0.169	94.9
					$G_1(x) = 2\log(1 + x/2)$ & $G_2(x) = \log(1 + x)$							
200	$(0, 0.5)$	β_1	-0.010	0.246	0.260	96.4	$(0.5, -0.5)$	β_1	0.012	0.269	0.270	95.1
		β_2	0.018	0.452	0.452	95.6		β_2	0.016	0.457	0.458	95.1
		γ	0.033	0.262	0.281	95.8		γ	0.024	0.264	0.266	95.2
400	$(0, 0.5)$	β_1	-0.009	0.175	0.176	95.1	$(0.5, -0.5)$	β_1	0.011	0.188	0.185	94.4
		β_2	0.011	0.311	0.317	95.2		β_2	-0.013	0.300	0.303	95.0
		γ	0.021	0.180	0.185	95.6		γ	0.016	0.184	0.188	95.2
					$G_1(x) = \log(1 + x)$ & $G_2(x) = \log(1 + x)$							
200	$(0, 0.5)$	β_1	0.005	0.279	0.280	95.3	$(0.5, -0.5)$	β_1	-0.010	0.276	0.287	95.0
		β_2	0.022	0.486	0.490	96.2		β_2	-0.023	0.486	0.485	95.2
		γ	0.028	0.275	0.299	96.1		γ	0.034	0.255	0.256	94.6
400	$(0, 0.5)$	β_1	-0.003	0.188	0.187	95.6	$(0.5, -0.5)$	β_1	0.009	0.196	0.196	95.0
		β_2	0.001	0.341	0.342	94.0		β_2	0.014	0.346	0.347	95.1
		γ	0.026	0.197	0.200	94.8		γ	0.023	0.191	0.190	94.8

Table 2. Estimation of regression parameters with the misspecified latent variable distribution

(β_{10}, β_{20})	Par	Bias	SSE	SEE	CP	(β_{10}, β_{20})	Par	Bias	SSE	SEE	CP
$G_1(x) = x$ & $G_2(x) = x$											
$(0, 0.5)$	β_1	-0.015	0.224	0.226	95.6	$(0.5, -0.5)$	β_1	0.035	0.253	0.252	95.8
	β_2	-0.016	0.396	0.395	95.2		β_2	0.031	0.436	0.428	94.8
$G_1(x) = 2\log(1 + x/2)$ & $G_2(x) = 2\log(1 + x/2)$											
$(0, 0.5)$	β_1	-0.006	0.280	0.278	95.2	$(0.5, -0.5)$	β_1	0.028	0.288	0.290	94.8
	β_2	0.015	0.501	0.522	96.8		β_2	-0.019	0.476	0.480	96.8
$G_1(x) = 2\log(1 + x/2)$ & $G_2(x) = \log(1 + x)$											
$(0, 0.5)$	β_1	0.014	0.281	0.286	95.2	$(0.5, -0.5)$	β_1	0.020	0.304	0.304	96.0
	β_2	0.022	0.525	0.541	96.6		β_2	-0.018	0.476	0.497	96.0
$G_1(x) = \log(1 + x)$ & $G_2(x) = \log(1 + x)$											
$(0, 0.5)$	β_1	0.011	0.282	0.297	95.4	$(0.5, -0.5)$	β_1	0.019	0.304	0.309	95.0
	β_2	0.022	0.562	0.569	96.2		β_2	-0.025	0.512	0.517	94.8

Table 3. Comparison of the proposed estimator and that given in Wang et al. (2015)

			Proposed method		Wang et al.(2015)	
γ_0	$(\beta_0^{(1)}, \beta_0^{(2)})$	Par	Bias	SSE	Bias	SSE
0.5	(0, 0.5)	$\beta^{(1)}$	-0.003	0.250	-0.002	0.250
		$\beta^{(2)}$	0.022	0.256	0.025	0.257
		γ	0.009	0.215	0.026	0.232
	(0.5, -0.5)	$\beta^{(1)}$	0.004	0.263	0.006	0.263
		$\beta^{(2)}$	-0.026	0.278	-0.025	0.277
		γ	0.022	0.236	0.042	0.253
1	(0, 0.5)	$\beta^{(1)}$	0.001	0.300	-0.002	0.299
		$\beta^{(2)}$	0.027	0.310	0.029	0.313
		γ	0.035	0.292	0.045	0.328
	(0.5, -0.5)	$\beta^{(1)}$	-0.009	0.320	-0.011	0.322
		$\beta^{(2)}$	0.018	0.291	0.019	0.293
		γ	0.042	0.397	0.040	0.420
1.5	(0, 0.5)	$\beta^{(1)}$	0.004	0.343	0.005	0.345
		$\beta^{(2)}$	0.017	0.355	0.019	0.354
		γ	0.032	0.459	0.100	0.472
	(0.5,-0.5)	$\beta^{(1)}$	0.003	0.351	0.001	0.343
		$\beta^{(2)}$	-0.029	0.345	-0.035	0.351
		γ	0.024	0.392	0.083	0.461

Table 4. Results on the analysis of the chlamydia and gonorrhea data assuming the same covariate effects

Frailty distribution		PH frailty model			PO frailty model			Optimal model		
		Est	Std	p-value	Est	Std	p-value	Est	Std	p-value
Log-normal	Gender	0.097	0.096	0.311	0.107	0.111	0.334	0.119	0.085	0.165
	C-America	-0.749	0.096	<0.001	-0.786	0.094	<0.001	-0.826	0.105	<0.001
	Symptoms	0.507	0.121	<0.001	0.549	0.144	<0.001	0.593	0.123	<0.001
Gamma	Gender	0.101	0.104	0.332	0.112	0.107	0.293	0.115	0.083	0.166
	C-America	-0.756	0.099	<0.001	-0.797	0.101	<0.001	-0.810	0.097	<0.001
	Symptoms	0.513	0.128	<0.001	0.560	0.127	<0.001	0.574	0.147	<0.001

Table 5. Results on the analysis of the chlamydia and gonorrhea data assuming different covariate effects

Frailty distribution		PH frailty model			PO frailty model			Optimal model		
		Est	Std	p-value	Est	Std	p-value	Est	Std	p-value
Log-normal						Chlamydia				
	Gender	0.060	0.117	0.611	0.065	0.088	0.460	0.073	0.117	0.532
	C-America	-0.619	0.108	<0.001	-0.650	0.097	<0.001	-0.692	0.113	<0.001
	Symptoms	0.296	0.125	0.017	0.314	0.111	0.005	0.340	0.148	0.021
						Gonorrhea				
	Gender	0.318	0.189	0.091	0.335	0.190	0.078	0.363	0.206	0.079
	C-America	-1.577	0.303	<0.001	-1.599	0.346	<0.001	-1.636	0.298	<0.001
	Symptoms	0.500	0.170	<0.001	1.346	0.203	<0.001	1.402	0.243	<0.001
Gamma						Chlamydia				
	Gender	0.063	0.088	0.476	0.062	0.093	0.505	0.061	0.101	0.545
	C-America	-0.626	0.104	<0.001	-0.622	0.112	<0.001	-0.618	0.082	<0.001
	Symptoms	0.301	0.130	0.020	0.296	0.130	0.023	0.293	0.142	0.038
						Gonorrhea				
	Gender	0.328	0.223	0.143	0.325	0.214	0.129	0.323	0.217	0.137
	C-America	-1.582	0.261	<0.001	-1.578	0.266	<0.001	-1.574	0.304	<0.001
	Symptoms	1.322	0.217	<0.001	1.316	0.217	<0.001	1.313	0.195	<0.001