Biomedical Statistics and Informatics
2024, Vol. 9, No. 3, pp. 39-57
https://doi.org/10.11648/j.bsi.20240903.11

Research Article

Penalized Likelihood Estimation for Current Status Data with Informative Censoring

Alhassan Faisal[1, *], Samuel Iddi[2], Ezekiel Nii Noye Nortey[2], Kwabena Doku-Amponsah[2]

[1]Department of Statistics, Faculty of Physical Sciences, University for Development Studies, Tamale, Ghana

[2]Department of Statistics and Actuarial Science, School of Physical and Mathematical Sciences, University of Ghana, Legon, Accra

Abstract

Current status data occurs when failure time of subjects in a survival study is only known to be either less or greater than the censoring time. Thus, the failure time is either left – or right – censored. Analyzing data of this structure under the Cox Proportional Hazards model with dependent censoring assumption can be challenging. To address this, a Penalized Maximum Likelihood Estimation (PMLE) approach was proposed. The unknown baseline cumulative hazard functions for both the failure time and the censoring time were estimated using splines. The advantage of penalized approach over unpenalized method is that that the desired smoothness level of the functions are controlled by their respective penalty terms. The possible dependence between the failure and censoring times was accounted for using the gamma-frailty model. An easy to implement hybrid computational algorithm is proposed to estimate the PMLEs and the Bayesian technique was employed for the estimation of the variances of the parameters. Extensive simulation studies were conducted to assess the statistical properties of the PMLEs. It was observed that the realized estimators were not only consistent, asymptotically normal and efficient, but also, were robust to the number of knots chosen, the proportion of dependent censoring used and the frailty distribution employed. The proposed PMLE method was further applied to real data obtained from tumorigenicity experiment.

Keywords

Current Status Data, Splines, Proportional Hazards Model, Penalized Maximum Likelihood Estimation, Frailty Model

1. Introduction

Current status data also known as interval-censored case I occurs when the failure time T_i is not exactly observed but only known to be less or greater than the examination time C_i. Thus, T_i is either left-censored or right-censored at C_i. This type of data is encountered in many areas of studies especially in animal carcinogenicity experiments (see [10]), where animals are examined for the presence or absence of the event under study at their death or sacrifice times C_i after exposing them to cancerous environment. If C_i and T_i are assumed to be independent, then the censoring is said to be noninformative otherwise, it is informative. Many authors have proposed methods of analyzing noninformative current status data (see

*Corresponding author: afaisal@uds.edu.gh (Alhassan Faisal)

Received: 23 March 2024; **Accepted:** 26 August 2024; **Published:** 29 October 2024

[4, 8, 17]). However, the assumption of noninformative censoring may be violated in some circumstances and therefore, analyzing such data using methods developed for noninformative censoring can lead to bias inferences. A number of authors have also developed procedures for the analysis of informative current status data (see [7, 10]). However, these authors did so in a restrictive context of animal tumorigenicity experiments, where observation time is either at death or random sacrifice times. If an animal dies naturally, this death may have something to do with both the onset of the tumor and the treatment it is exposed to. These authors estimated the baseline cumulative hazard functions (CHFs) using transition functions and adopted expectation maximization (EM) algorithm for estimation after using the piecewise constant modeling approach to approximate those transition functions. These procedures are not only computationally cumbersome but also provides a highly restrictive application in practice. The concept of frailty (see [15]) has been the most commonly used approach in literature to account for the informative censoring with current status data.

For unrestrictive procedure, Zhang et al. (see [19]) proposed a novel method by cleverly merging the hazards for T_i, assumed to follow additive hazards frailty model and the hazards for C_i, assumed to follow proportional hazards frailty model, into a form of an additive hazards model. This approach though easy to implement, estimates only the regression parameters and avoids the estimation of the baseline hazard function as well as the frailty parameter. Chen et al. (see [1]) also proposed a class of semiparametric models and proposed a novel EM algorithm for the estimation, but approximated the baseline functions using piecewise constant approximation. Lu and Li (see [8]) also discussed sieve EM approach for the fitting but used splines to approximate the baseline cumulative hazard function (CHF) for the T_i and piece-wise constant for C_i. The major limitation of EM based on piece-wise constant approach is that, it does not produce the required smooth curves for a clearer trend of how the hazard of the event changes with time. In addition, if the piecewise constant assumption of the baseline CHFs is not the case in reality, then the accuracy of the corresponding inferences will be in doubt. Moreover, the EM based on sieve methods of estimation is not only complex in its approach, but also often has slow convergence if the initial values are not correctly chosen. Furthermore, separate computation is needed for the variance estimates. These difficulties motivated this study. Thus, in this work, an extension of the arguments of the authors (see [1, 9, 11]) to the environment of penalized likelihood estimation is made. The main goal of this paper therefore, is to propose a Penalized Maximum Likelihood Estimation (PMLE) procedure for the Proportional Hazard model under informative current status data, where the baseline CHFs for the T_i and C_i are approximated using splines (see [14]) and the informative censoring is accounted for using gamma frailty variable.

The paper is organized as follows: Section 2 presents the Materials and Methods whiles Section 3 contains the Results and Discussions whiles section 4 contains the concluding remarks.

2. Materials and Methods

2.1. Notations, Models and Penalized Log-Likelihood Function Construction

Suppose in a survival study, there exists one of two potential observation times for each individual: either at a pre-specified time C_i^* or random time C_i to ascertain whether or not the event of interest has occurred. Thus, the observation time is $\tilde{C}_i = min(C_i, C_i^*)$. If $\tilde{C}_i = C_i$, it is considered informative with an indicator $\Delta_i = 1$ and survival function S_C and if $\tilde{C}_i = C_i^*$, it is considered noninformative with indicator $\Delta_i = 0$. For the failure times T_i with the survival function S_T, we define the indicator $\delta_i = 1$ if $T_i \leq \tilde{C}_i$ and $\delta_i = 0$ if otherwise. If Z denotes time-independent p-covariates and the possible correlation between T_i and $\tilde{C}_i$ is accounted for by the individual's frailty b_i. Then the current status data consist of the set $X_i = (\tilde{C}_i, \delta_i, \Delta_i, Z_i, b_i, i = 1,2, \dots, n)$. Therefore, evaluating the effects of Z on T_i and on C_i using the Cox frailty model can be expressed as;

$$\Lambda_T(t_i|Z_i, b_i) = b_i \Lambda_{0t}(t_i)\, exp(\beta^T Z_i) \quad (1)$$

$$\Lambda_C(c_i|Z_i, b_i) = b_i \Lambda_{0c}(c_i)\, exp(\gamma^T Z_i) \quad (2)$$

where $\Lambda_{0t}(t_i)$ and $\Lambda_{0c}(c_i)$ are unspecified baseline CHFs of T_i and C_i respectively; β and γ denote the corresponding regression parameters of Z. The b_i can obey any distribution with strictly positive support. However, the gamma distribution has been widely used for modeling dependence in multivariate survival time (see [5]), because of the possibility of obtaining a closed form expression of marginal likelihood functions using Laplace transform (see [6]) defined as

$$L_b(s) = E(e^{-sb}) = \int_0^\infty e^{-sb} g(b;\theta) db_i = (1 + \theta s)^{-\frac{1}{\theta}} \quad (3)$$

where $g(b_i;\theta) = \frac{b_i^{\left(\frac{1}{\theta}-1\right)} exp\left(-\frac{b_i}{\theta}\right)}{\theta^{\frac{1}{\theta}} \Gamma\left(\frac{1}{\theta}\right)}$ is the gamma distribution with mean 1 and variance θ. High value of θ indicates a stronger correlation between the T_i and $\tilde{C}_i$ and vice-versa. It can be seen that the likelihood function for the complete current status dataset would consist of four cases depending on the categories defined by Δ_i and δ_i. If T_i and $\tilde{C}_i$ are assumed to be independent only when b_i and Z_i are given, then the joint conditional likelihood function associated with these categories labeled A to D are as follows;

(A) This category comprises of individuals with right-censored failure times at non-informative observation times

$$L_1(\theta_1,\theta_2|Z_i,b_i) = P(\delta_i = 0,\Delta_i = 0|Z_i,b_i) = P\big(T_i > \tilde{C}_i|Z_i,b_i\big)P\big(C_i > \tilde{C}_i|Z_i,b_i\big) = S_T(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i)$$

$$= exp(-b_i\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i))\,exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)) = exp\{-b_i(\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i) + \Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\} \quad (4)$$

Integrating out b_i in (4) using the concept in (3), we obtain the unconditional likelihood as

$$L_1 = L_1(\Phi|Z_i) = \int_0^\infty exp\{-b_i(\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i) + \Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}\,g(b_i,\theta)\,db_i$$

$$= \{1 + \theta(\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i) + \Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}^{-\frac{1}{\theta}} \quad (5)$$

where $\Phi = (\theta_1,\theta_2,\theta)$ with $\theta_1 = (\beta,\Lambda_{0t}(\cdot))$ and $\theta_2 = (\gamma,\Lambda_{0c}(\cdot))$ are the parameters to be estimated.

(B) This category comprises of individuals with left-censored failure times at non-informative observation times

$$L_2(\theta_1,\theta_2|Z_i,b_i) = P(\delta_i = 1,\Delta_i = 0|Z_i,b_i) = P\big(T_i \le \tilde{C}_i|Z_i,b_i\big)P\big(C_i > \tilde{C}_i|Z_i,b_i\big)$$

$$= F_T(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i) = [1 - S_T(\tilde{c}_i|Z_i,b_i)]S_C(\tilde{c}_i|Z_i,b_i) = S_C(\tilde{c}_i|Z_i,b_i) - S_T(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i)$$

The unconditional joint likelihood can be obtained by;

$$L_2 = L_2(\Phi|Z_i) = \int_0^\infty S_C(\tilde{c}_i|Z_i,b_i)g(b_i,\theta)db_i - \int_0^\infty S_T(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i)\,g(b_i,\theta)\,db_i$$

$$= \Big[\int_0^\infty exp\{-b_i(\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}\,g(b_i,\theta)\,db_i\Big] - L_1 = (1 + \theta(\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)))^{-\frac{1}{\theta}} - L_1 \quad (6)$$

(C) This category comprises of individuals with right-censored failure times at the informative observation times

$$L_3(\theta_1,\theta_2|Z_i,b_i) = P(\delta_i = 0,\Delta_i = 1|Z_i,b_i) = P\big(T_i > \tilde{C}_i|Z_i,b_i\big)P\big(C_i = \tilde{C}_i|Z_i,b_i\big)$$

$$= S_T(\tilde{c}_i|Z_i,b_i)\,f_C(\tilde{c}_i|Z_i,b_i) = S_T(\tilde{c}_i|Z_i,b_i)\,\lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)$$

The unconditional likelihood is therefore

$$L_3 = L_3(\Phi|Z_i) = \int_0^\infty S_T(\tilde{c}_i|Z_i,b_i)\,\lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)\,g(b_i,\theta)db_i$$

$$= \int_0^\infty \lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)\,S_T(\tilde{c}_i|Z_i,b_i)g(b_i,\theta)db_i$$

$$= \int_0^\infty b_i\,\lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)\,exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\,exp(-b_i\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i))\,g(b_i;\theta)db_i$$

$$= -\int_0^\infty \frac{d}{d\Lambda_{0c}}\{exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}\,exp(-b_i\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i))\,g(b_i;\theta)db_i$$

$$= -\frac{d}{d\Lambda_{0c}}\Big[\int_0^\infty exp\big(-b_i(\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i) + \Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i))\big)\,g(b_i;\theta)db_i\Big]$$

$$= -\frac{d}{d\Lambda_{0c}}\Big[\big(1 + \theta(\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i) + \Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\big)^{-\frac{1}{\theta}}\Big]$$

$$= \lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)\big(1 + \theta(\Lambda_{0t}(\tilde{c}_i)\,exp(\beta^T Z_i) + \Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\big)^{-\frac{1}{\theta}-1} \quad (7)$$

(D) This category comprises of individuals with left-censored failure times at the informative observation times

$$L_4(\theta_1,\theta_2|Z_i,b_i) = P(\delta_i = 1,\Delta_i = 1|Z_i,b_i) = P\big(T_i \le \tilde{C}_i|Z_i,b_i\big)P\big(C_i = \tilde{C}_i|Z_i,b_i\big) = F_T(\tilde{c}_i|Z_i,b_i)\,f_C(\tilde{c}_i|Z_i,b_i) =$$
$$F_T(\tilde{c}_i|Z_i,b_i)\,\lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i) = [1 - S_T(\tilde{c}_i|Z_i,b_i)]\,\lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)$$

$$= \lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i) - S_T(\tilde{c}_i|Z_i,b_i)\lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)$$

The unconditional likelihood is therefore

$$L_4 = L_4(\Phi|Z_i) = \int_0^\infty \lambda_C(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i)g(b_i;\theta)db - \int_0^\infty S_T(\tilde{c}_i|Z_i,b_i)\lambda_C(\tilde{c}_i|Z_i,b_i)S_C(\tilde{c}_i|Z_i,b_i)g(b_i;\theta)db_i$$

$$= \left[\int_0^\infty \lambda_C(\tilde{c}_i|Z_i,b_i)\,S_C(\tilde{c}_i|Z_i,b_i)\,g(b_i;\theta)db\right] - L_3$$

$$= \left[\int_0^\infty b_i\lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)\,exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\,g(b_i;\theta)db_i\right] - L_3$$

$$= \left[-\int_0^\infty \frac{d}{d\Lambda_{0c}}\{exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}g(b_i;\theta)db_i\right] - L_3$$

$$= -\frac{d}{d\Lambda_{0c}}\left[\int_0^\infty\{exp(-b_i\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))\}g(b_i;\theta)db_i\right] - L_3$$

$$= -\frac{d}{d\Lambda_{0c}}\left[(1+\theta\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))^{-\frac{1}{\theta}}\right] - L_3$$

$$= \lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i)(1+\theta\Lambda_{0c}(\tilde{c}_i)\,exp(\gamma^T Z_i))^{-\frac{1}{\theta}-1} - L_3 \tag{8}$$

he full likelihood and its log-likelihood function therefore can be written respectively as

$$L_f(\Phi) = \prod_{i=1}^n\left[(L_1)^{(1-\delta_i)(1-\Delta_i)}(L_2)^{\delta_i(1-\Delta_i)}(L_3)^{(1-\delta_i)\Delta_i}(L_4)^{\delta_i\Delta_i}\right] \tag{9}$$

$$l(\Phi) = \sum_{i=1}^n[(1-\delta_i)(1-\Delta_i)\,log\,L_1 + \delta_i(1-\Delta_i)\,log\,L_2 + (1-\delta_i)\Delta_i\,log\,L_3 + \delta_i\Delta_i\,log\,L_4] \tag{10}$$

In reality, Λ_{0t} and Λ_{0c} are expected to be smooth so that its trend of changing hazard over time can be clearly observed. Therefore, to control their possible roughness, each was penalized by subtracting their respective penalty terms from the function in (10) to obtain

$$pl(\Phi) = l(\Phi) - \frac{\kappa_1}{2}\int \Lambda_{0t}''^{\,2}(t)dt - \frac{\kappa_2}{2}\int \Lambda_{0c}''^{\,2}(t)dt \tag{11}$$

where $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$ are the smoothing parameters that control the smoothness level of $\Lambda_{0t}(t)$ and $\Lambda_{0c}(t)$ respectively. Maximizing (11) for given κ_1 and κ_2 gives us the PMLE Φ.

2.2. Modeling Baseline CHFs $\Lambda_{0t}(\cdot)$ and $\Lambda_{0c}(\cdot)$ With Splines

The $\Lambda_{0t}(\cdot)$ and $\Lambda_{0c}(\cdot)$ were estimated using the integrated spline functions. This approach does not only allow flexible shapes of the function but also ensures a significant reduction in the number of required parameters to estimate. One can refer to [2] for details on splines. Thus, each baseline function is approximated as (see [17]);

$$\left.\begin{aligned}\hat{\Lambda}_{0t}(t) &\approx \sum_{j=1}^m \hat{\alpha}_j I_j(t)\\ \hat{\Lambda}_{0c}(t) &\approx \sum_{j=1}^m \hat{\eta}_j I_j(t)\end{aligned}\right\} \tag{12}$$

where α_j and η_j are the spline coefficients that are constraint to be positive using the square transformation; $m = q_n + k$ is the required number of spline basis functions; q_n is the number of knots and $I_j(t)$ is the integrated spline basis of degree k defined as (see [14]);

$$I_j(x;k) = \int_0^x M(u,k)du \tag{13}$$

where $M(x,k)$ is the monotone spline calculated recursively by the formula as;

$$M_j(x;k) = \begin{cases} \frac{k[(x-t_j)M_J(x;k-1)+(t_{j+k}-x)M_{J+1}(x;k-1)]}{(k-1)(t_{j+k}-t_j)}, & t_j \le x < t_{j+k} \\ 0, & elsewhere \end{cases} \quad (14)$$

$$\text{with } M_j(x;1) = \begin{cases} \frac{1}{(t_{j+k}-t_j)}, & t_j \le x < t_{j+1} \\ 0, & elsewhere \end{cases} \quad (15)$$

where $t_1, t_2, \ldots, t_m$ is a sequence of increasing knots in the interval $[\tilde{C}_{min}, \tilde{C}_{max}]$. In this study, q_n is determined using the relation $q_n = \left\lfloor n^{\frac{1}{3}} \right\rfloor$, where $\left\lfloor n^{\frac{1}{3}} \right\rfloor$ is the integer part of $n^{\frac{1}{3}}$ and then position the knots using the quantile method. These spline basis functions can easily be computed using the R package '*splines2*' developed by [18].

To further reduce the computational burden of (11), the roughness penalty terms for $\Lambda_{0t}(\cdot)$ and $\Lambda_{0c}(\cdot)$ was approximated in terms of splines using the argument of [3] as;

$$\frac{\kappa_1}{2}\int \Lambda''^2_{0t}(t)dt \approx \frac{\kappa_1}{2}\sum_{j=3}^{m}(\Delta^2\alpha_j)^2 = \frac{\kappa_1}{2}\sum_{j=3}^{m}(\alpha_j - 2\alpha_{j-1} + \alpha_{j-2})^2 = \frac{\kappa_1}{2}\alpha^T A^T A\alpha = \frac{\kappa_1}{2}\alpha^T R_1 \alpha \quad (16)$$

$$\frac{\kappa_2}{2}\int \Lambda''^2_{0c}(t)dt \approx \frac{\kappa_2}{2}\sum_{j=3}^{m}(\Delta^2\eta_j)^2 = \frac{\kappa_2}{2}\sum_{j=3}^{m}(\eta_j - 2\eta_{j-1} + \eta_{j-2})^2 = \frac{\kappa_2}{2}\eta^T C^T C\eta = \frac{\kappa_2}{2}\eta^T R_2 \eta \quad (17)$$

where $A = C$ are suitably defined penalty matrices each of size $(m-2)\times m$ based on the second-order difference operator Δ^2. Therefore, (11) can be re-written as;

$$pl(\Phi) = l(\Phi) - \frac{\kappa_1}{2}\alpha^T R_1 \alpha - \frac{\kappa_2}{2}\eta^T R_2 \eta \quad (18)$$

Given κ_1 and κ_2, we can obtain the PMLE, $\hat{\Phi} = (\hat{\beta}, \hat{\gamma}, \hat{\alpha}, \hat{\eta}, \hat{\theta})^T$ by maximizing (18).

2.3. Selection of the Smoothing Parameters κ_1 and κ_2

O'Sullivan (see [13]) proposed an approximated cross validation score (CVS) for the determination of optimal κ_1 and κ_2, which is herein adapted, and is given by

$$CVS(\kappa_h) = -\frac{1}{n}l(\hat{\Phi}_{\kappa_h}) + \frac{1}{n}trace\left[\left(\hat{H}_{pl}(\hat{\Phi}_{\kappa_h})\right)^{-1}\hat{H}_l(\hat{\Phi}_{\kappa_h})\right] \quad (19)$$

where $l(\hat{\Phi}_{\kappa_h})$ is defined in (10), $\hat{H}_{pl}(\hat{\Phi}_{\kappa_h})$ and $\hat{H}_l(\hat{\Phi}_{\kappa_h})$ are respectively the Hessian matrices of (18) and (10). A search for the value of κ_h, which maximizes (19) using the grid search is then carried out. But simultaneous determination of κ_1 and κ_2 may pose computational challenges. For that matter, a proposal to maximize their respective penalized log-likelihood under the assumption of noninformative observation times was made. Thus, to choose κ_1, the data $\left(\tilde{C}_i, \delta_i = I(T_i \le \tilde{C}_i), Z_i\right)$ is used and the log-likelihood and penalized functions can be written respectively as;

$$l(\theta_1) = \sum_{i=1}^{n}\left(\delta_i\, log\left(1 - exp\left(-\sum_{j=1}^{m}\alpha_j I_j(t)\, exp(\beta^T Z_i)\right)\right) - (1-\delta_i)\sum_{j=1}^{m}\alpha_j I_j(t)\, exp(\beta^T Z_i)\right) \quad (20)$$

$$pl(\theta_1) = l(\theta_1) - \frac{\kappa_1}{2}\alpha^T R_1 \alpha \quad (21)$$

where $\theta_1 = (\beta_1, \ldots, \beta_p, \alpha_1, \ldots, \alpha_m)^T$.

The approximate CVS for (21) is

$$CVS(\kappa_1) = -\frac{1}{n}l(\hat{\theta}_{\kappa_1}) + \frac{1}{n}trace\left[\left(\hat{H}_{pl}(\hat{\theta}_{\kappa_1})\right)^{-1}\hat{H}_l(\hat{\theta}_{\kappa_1})\right] \quad (22)$$

where $l(\cdot)$ is defined in (20), $\hat{H}_{pl}(\hat{\theta}_{\kappa_1})$ and $\hat{H}_l(\hat{\theta}_{\kappa_1})$ are the Hessian matrices of (21) and (20) computed at $\hat{\theta}_{\kappa_1}$, R_1 is as defined previously. Then (22) is optimized using grid search to obtain κ_1 and a good initial guess of the vector $\hat{\theta}_1$. Similarly, to choose κ_2, we treated the data set $\left(\tilde{C}_i, \Delta_i = I(C_i \le C_i^*), Z_i\right)$ as right-censored data where $\tilde{C}_i = min(C_i, C_i^*)$ with indicator function $\Delta_i = I(C_i \le C_i^*)$. Therefore, the appropriate log-likelihood and penalized functions respectively;

$$l(\theta_2) = \sum_{i=1}^{n}\left(\Delta_i\, log\left(\sum_{j=1}^{m}\eta_j M_j(t)\, exp(\gamma^T Z_i)\right) - \sum_{j=1}^{m}\eta_j I_j(t)\, exp(\gamma^T Z_i)\right) \quad (23)$$

$$pl(\theta_2) = l(\theta_2) - \frac{\kappa_2}{2}\eta^T R_2 \eta \quad (24)$$

where $\theta_2 = (\gamma_1, \ldots, \gamma_p, \eta_1, \ldots, \eta_m)^T$.

The approximate CVS of (24) can be written as;

$$CVS(\kappa_2) = -\frac{1}{n} l(\hat{\theta}_{\kappa_2}) + \frac{1}{n} trace\left[\left(\hat{H}_{pl}(\hat{\theta}_{\kappa_2})\right)^{-1} \hat{H}_l(\hat{\theta}_{\kappa_2})\right] \quad (25)$$

where $l(\cdot)$is defined in (23); $\hat{H}_{pl}(\hat{\theta}_{\kappa_2})$ and $\hat{H}_l(\hat{\theta}_{\kappa_2})$ respectively represent the Hessian matrix of (24) and (23) evaluated at $\hat{\theta}_{\kappa_2}$, R_2 is as defined previously. We then optimize (25) using the grid search to obtain κ_2 and a good initial guess of the vector $\hat{\theta}_2$. In this study, 50 grid values of κ_h, $h = 1{,}2$ was chosen, equally spaced between 0.001 and 1.5 in both the simulation and practical data analyses.

2.4. Proposed Computational Algorithm

To maximize (19), the following computing algorithm is proposed;

Step 1: Choose a set of 50 grid points for τ, equally spaced from 0.001 to 1.5 and let $i = 1$

Step 2: With τ_i, obtain the PMLEs $\hat{\theta}_{\tau_i,}$ of equation (22) using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) Optim function in R. Refer to [12] for details.

Step 3: Compute the CVS, $CVS(\tau_i)$ using equation (22), let $i = i + 1$

Step 4: Iterate between steps 2 and 3 until all the 50 grid points are exhausted move to step 5

Step 5: Select the value τ that resulted in largest value of $CVS(\tau)$ from step 4

Step 6: Compute the optimal κ_1 parameter using the expression $\kappa_1 = \tau \times n^{-\frac{1}{3}}$

Step 7: Recall step 1 and with τ_i, obtain the PMLEs $\hat{\theta}_{\tau_i}$ of equation (24) using the BFGS optim in R

Step 8: Compute the CVS, $CVS(\tau_i)$ using equation (25), let $i = i + 1$

Step 9: iterate between steps 7 and 8 until all the 50 grid points are exhausted and move to step 10

Step 10: Select the value τ that resulted in largest value of $CVS(\tau)$ from step 9

Step 11: Compute the optimal κ_2 parameter using the expression $\kappa_2 = \tau \times n^{-\frac{1}{3}}$

Step 12: Using the initial values $\hat{\Phi}^{(0)} = (\hat{\theta}_{\kappa_1}, \hat{\theta}_{\kappa_2}, \hat{\theta})^T$, where $\hat{\theta}_{\kappa_1}$and $\hat{\theta}_{\kappa_2}$ are as obtained in steps 5 and 10 respectively and $\hat{\theta} = 1$. Maximize (18) again using the BFGS optim in R to obtain the final PMLEs, $\hat{\Phi} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta})^T$.

Step 13: obtain the estimates of $\Lambda_{0t}(t)$ and $\Lambda_{0c}(t)$ and using equation (12) appropriately.

2.5. Asymptotic Properties of the PMLEs $\hat{\Phi}$

To assess the asymptotic properties of the PMLEs, we employed the Bayesian technique proposed by Wahba (1983). Thus, Φ is treated as random variable with $\kappa_h J_h(\Phi)$ as its prior log-likelihood and $pl(\Phi)$ as the posterior log-likelihood. Since the posterior distribution is conditional upon observing the sample, its distribution is used to make statements about this random quantity Φ

Theorem: If conditions (A1) – (A8) below are met, then as $n \to \infty$, the PMLE $\hat{\Phi}$ asymptotically follows a multivariate posterior Gaussian distribution with mean Φ_0 and variance – covariance matrix as $Var(\hat{\Phi}) = I(\Phi_0)^{-1}$. That is

$\sqrt{n}(\hat{\Phi}_n - \Phi_0) \to N(0, (I(\Phi_0))^{-1})$ where $I(\Phi_0)$ is the efficient information for Φ.

(A1) the observed data $A = \{y_i : p(y_i|\Phi); i = 1{,}2, \ldots, n\}$, are independently and identically distributed and does not depend on Φ.

(A2) The censoring time and the true unobserved failure time are conditionally independent given the covariates and frailties.

(A3) $\lim_{n\to\infty}(n^{-1}\sum_{i=1}^{n} l_i(\Phi))$ exists and has a unique maximum at $\Phi = \Phi_0$, which belongs to the interior of the compact (parameter) space Θ.

(A4) $\mu_n = o(n^{-\frac{1}{2}})$ and $\omega_n = o(n^{-\frac{1}{2}})$

(A5) There exists a measurable function $g(y_i) = g(\tilde{C}_i, \delta_i, \Delta_i, Z_i)$ with $E(g(y_i)) < \infty$ which satisfies $|log\, p\,(y_i|\Phi)| \leq g(y_i)$ for all $\Phi \in \Theta$.

(A6) $l(\Phi)$ is two times continuously differentiable in a neighborhood of Φ_0 such that $E_{\Phi_0}\left(l_i'(\Phi)\right) = 0$ and $Var_{\Phi_0}\left(l_i'(\Phi)\right) = E_{\Phi_0}\left(l_i'(\Phi)\right)^2 = E_{\Phi_0}\left(l_i'(\Phi) l_i'(\Phi)^T\right) = -E_{\Phi_0}\left(l_i''(\Phi)\right) = I(\Phi_0)$

(A7) The penalty function $J_t(\Phi)$ is continuous and bounded over Θ where both derivatives $J_t'(\Phi)$ and $J_t''(\Phi)$ exist for all $\Phi \in \Theta$, and $J_t''(\Phi)$ is bounded in a neighborhood of Φ_0.

(A8) The penalty function $J_c(\Phi)$ is continuous and bounded over Θ and both derivatives $J_c'(\Phi)$ and $J_c''(\Phi)$ exist for all $\Phi \in \Theta$, and $J_c''(\Phi)$ is bounded in a neighborhood of Φ_0.

Proof of the Theorem

In order to prove this Bayesian estimator, the asymptotic proof was restricted to the assumption that the number of basis functions are fixed but large and with fixed κ_1 and κ_2. The parameters $\Phi = (\beta', \gamma', \alpha', \eta', \theta)$ has a length of $r = 2p + 2m + 1$. If we let $\Phi_0 = (\beta_0', \gamma_0', \alpha_0', \eta_0', \theta_0)$ be the true value of Φ and denote the observed data for the i^{th} individual as $y_i = (\tilde{C}_i, \delta_i, \Delta_i, Z_i)$ with $\tilde{C}_i = min(C_i, C_i^*)$, $\Delta_i = I(\tilde{C}_i = C_i)$ and $\delta_i = I(T_i \leq \tilde{C}_i)$. Given the b_i for this i^{th} individual at the observation time $\tilde{C}_i = \tilde{c}_i$, then the unconditional joint density function can be written as;

$$p(y_i|\Phi) = \int_0^{\infty} \begin{bmatrix} \{[F_T(\tilde{c}_i)f_C(\tilde{c}_i)]^{\delta_i}[S_T(\tilde{c}_i)f_C(\tilde{c}_i)]^{(1-\delta_i)}\}^{\Delta_i} \\ \times \{[F_T(\tilde{c}_i)S_C(\tilde{c}_i)]^{\delta_i}[S_T(\tilde{c}_i)S_C(\tilde{c}_i)]^{(1-\delta_i)}\}^{(1-\Delta_i)} g(b_i,\theta)db_i \end{bmatrix}. \quad (26)$$

where $F_T(\tilde{c}_i)$ and $S_T(\tilde{c}_i)$ represent the cumulative distribution and survival functions of T_i given Z_i and b_i respectively, whiles $S_C(\tilde{c}_i)$ and $f_C(\tilde{c}_i)$ are respectively survival and density functions of $\tilde{C}_i$ given Z_i and b_i. The sample log-likelihood and penalized functions can be written as;

$$l(\Phi) = log\, L\,(\Phi) = log(\prod_{i=1}^{n} p(y_i|\Phi)) = \sum_{i=1}^{n} log\, p\,(y_i|\Phi) = \sum_{i=1}^{n} l_i(\Phi) \quad (27)$$

$$pl(\Phi) = l(\Phi) - \frac{\kappa_1}{2}J_t(\Phi) - \frac{\kappa_2}{2}J_c(\Phi) = \sum_{i=1}^{n} l_i(\Phi) - \frac{\kappa_1}{2}J_t(\Phi) - \frac{\kappa_2}{2}J_c(\Phi) \quad (28)$$

where $J_t(\Phi) = J_t(\alpha)$ and $J_c(\Phi) = J_c(\eta)$. As per A6, A7 and A8, let the first two derivatives of (28) be $pl'(\Phi)$ and $pl''(\Phi)$. If $\hat{\Phi}$ is the PMLE for Φ, then it is a fact that $pl'(\hat{\Phi}) = 0$. The second-order Taylor series expansion of (28) around the true Φ_0 gives

$$0 = pl'(\hat{\Phi}) = pl'(\Phi_0) + (\hat{\Phi}_n - \Phi_0)pl''(\Phi_n^*) \Longrightarrow (\hat{\Phi}_n - \Phi_0) = -\frac{pl'(\Phi_0)}{pl''(\Phi_n^*)} \quad (29)$$

where Φ_n^* is a vector between $\hat{\Phi}_n$ and Φ_0. Multiplying both sides of (29) by $\sqrt{n}$, we get

$$\sqrt{n}(\hat{\Phi}_n - \Phi_0) = \sqrt{n}\left(-\frac{pl'(\Phi_0)}{pl''(\Phi_n^*)}\right) = -\frac{\frac{1}{\sqrt{n}}pl'(\Phi_0)}{\frac{1}{n}pl''(\Phi_n^*)} \quad (30)$$

Next, it should be shown that as $n \to \infty$, the numerator and the denominator of (30) respectively behave as:
(a) $\frac{1}{\sqrt{n}}pl'(\Phi_0) \xrightarrow[d]{} N(0, I(\Phi_0))$ and (b) $\frac{1}{n}pl''(\Phi_n^*) \xrightarrow[p]{} I(\Phi_0)$
To prove (a), it can be written that

$$\frac{1}{\sqrt{n}}pl'(\Phi_0) = \frac{1}{\sqrt{n}}\left[l'(\Phi) - \frac{\kappa_1}{2}J_t'(\Phi) - \frac{\kappa_2}{2}J_c'(\Phi)\right] = \frac{l'(\Phi)}{\sqrt{n}} - \frac{\kappa_1}{2\sqrt{n}}J_t'(\Phi) - \frac{\kappa_2}{2\sqrt{n}}J_c'(\Phi) = n^{-\frac{1}{2}}\sum_{i=1}^{n} l_i'(\Phi) - \frac{\mu_n}{2}J_t'(\Phi) - \frac{\omega_n}{2}J_c'(\Phi) \quad (31)$$

where $\mu_n = \frac{\kappa_1}{\sqrt{n}}$, $\omega_n = \frac{\kappa_2}{\sqrt{n}}$ and $\sum_{i=1}^{n} l_i'(\Phi)$ can be seen as adding n independent and identically distributed (iid) gradient vectors of (27) with $l_i'(\Phi) = \frac{\partial\, log\, p(y_i|\Phi)}{\partial \Phi}, i = 1,2,\ldots,n$. As for the second and third terms, they tend to zero as $n \to \infty$. Therefore, (31) can be re-written as;

$$\frac{1}{\sqrt{n}}pl'(\Phi_0) = n^{-\frac{1}{2}}\sum_{i=1}^{n} l_i'(\Phi) = \frac{\sum_{i=1}^{n} l_i'(\Phi)}{\sqrt{n}} \quad (32)$$

Therefore, under condition (A6), the expectation (mean) can be found as

$$E_{\Phi_0}\left(\frac{1}{\sqrt{n}}pl'(\Phi_0)\right) = E_{\Phi_0}\left(\frac{\sum_{i=1}^{n} l_i'(\Phi)}{\sqrt{n}}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} E_{\Phi_0}(l_i'(\Phi)) = 0$$

For the variance, it is found as follows;

$$Var_{\Phi_0}\left(\frac{1}{\sqrt{n}}pl'(\Phi_0)\right) = Var_{\Phi_0}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_i'(\Phi)\right) = \frac{1}{n}\sum_{i=1}^{n} Var_{\Phi_0}(l_i'(\Phi)) = \frac{1}{n}(nI(\Phi_0)) = I(\Phi_0)$$

Therefore, by the Central Limit Theorem (CLT), $\frac{1}{\sqrt{n}}pl'(\Phi_0) \xrightarrow[d]{} N(0, I(\Phi_0))$ and therefore (a) is proven. For the (b), it can be written that

$$\frac{1}{n}pl''(\Phi_n^*) = \frac{1}{n}\left[l''(\Phi) - \frac{\kappa_1}{2}J_t''(\Phi) - \frac{\kappa_2}{2}J_c''(\Phi)\right] = \frac{1}{n}l''(\Phi) - \frac{\kappa_1}{2n}J_t''(\Phi) - \frac{\kappa_2}{2n}J_c''(\Phi) = \frac{1}{n}\sum_{i=1}^{n} l_i''(\Phi) - \frac{\mu_n}{2}J_t''(\Phi) - \frac{\omega_n}{2}J_c''(\Phi) \quad (33)$$

Where $\mu_n = \frac{\kappa_1}{n}$, $\omega_n = \frac{\kappa_2}{n}$ and the $\sum_{i=1}^{n} l_i''(\Phi)$ can be seen as the sum of n i.i.d. Hessian matrix of the (27). Again, the se-

cond and third terms tend to zero as $n \to \infty$. Therefore, (33) can be re-written as;

$$\frac{1}{n}pl''(\Phi_n^*) = \frac{1}{n}\sum_{i=1}^{n} l_i''(\Phi) = n^{-1}\sum_{i=1}^{n}\frac{\partial^2 \log p(y_i|\Phi_n^*)}{\partial\Phi\partial\Phi'} \tag{34}$$

Applying the weak law of large numbers (WLLN), (34) can be rewritten as

$$\frac{1}{n}pl''(\Phi_n^*) = \frac{1}{n}\sum_{i=1}^{n} l_i''(\Phi) = n^{-1}\sum_{i=1}^{n}\frac{\partial^2 \log p(y_i|\Phi_n^*)}{\partial\Phi\partial\Phi'} \xrightarrow[p]{} I(\Phi_0)$$

Hence (b) is proven.

Finally, the asymptotic behavior of the distribution $\sqrt{n}(\hat{\Phi}_n - \Phi_0)$ can be seen as;

$$E\left(\sqrt{n}(\hat{\Phi}_n - \Phi_0)\right) = E\left(\frac{\frac{1}{\sqrt{n}}pl'(\Phi_0)}{\frac{1}{n}pl''(\Phi_n^*)}\right) = \left\{-\frac{1}{n}pl''(\Phi_n^*)\right\}^{-1} E\left(\frac{1}{\sqrt{n}}pl'(\Phi_0)\right) = 0$$

$$Var\left(\sqrt{n}(\hat{\Phi}_n - \Phi_0)\right) = Var\left(\frac{\frac{1}{\sqrt{n}}pl'(\Phi_0)}{\frac{1}{n}pl''(\Phi_n^*)}\right) = \left\{\frac{1}{n}pl''(\Phi_n^*)\right\}^{-1} Var\left(\frac{1}{\sqrt{n}}pl'(\Phi_0)\right)\left\{\frac{1}{n}pl''(\Phi_n^*)\right\}^{-1}$$
$$=\{I(\Phi_0)\}^{-1}I(\Phi_0)\{I(\Phi_0)\}^{-1}=\{I(\Phi_0)\}^{-1}$$

$\Rightarrow \sqrt{n}(\hat{\Phi}_n - \Phi_0) \xrightarrow[d]{} N\left(0,\ I(\Phi_0)^{-1}\right)$ as required in the stated theorem.

If $Var(\hat{\Phi}) = I(\Phi)^{-1} = \begin{bmatrix} \tilde{H}_{\beta\beta}^{-1} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \tilde{H}_{\gamma\gamma}^{-1} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \tilde{H}_{\alpha\alpha}^{-1} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \tilde{H}_{\eta\eta}^{-1} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \tilde{H}_{\theta\theta}^{-1} \end{bmatrix}$, then by [16] an approximate 95% credible interval for the parameters $\beta_j,\ \gamma_j,\ \theta,\ \Lambda_{0t}(t)$ and $\Lambda_{0c}(t)$ are estimated as follows;

$$\left.\begin{aligned} \hat{\beta}_j \pm 1.96SE(\hat{\beta}_j) &\Rightarrow \hat{\beta}_j \pm 1.96\sqrt{\tilde{H}_{\beta\beta,j}^{-1}} \\ \hat{\gamma}_j \pm 1.96SE(\hat{\gamma}_j) &\Rightarrow \hat{\gamma}_j \pm 1.96\sqrt{\tilde{H}_{\gamma\gamma,j}^{-1}} \\ \hat{\theta} \pm 1.96SE(\hat{\theta}) &\Rightarrow \hat{\theta} \pm 1.96\sqrt{\tilde{H}_{\theta\theta}^{-1}} \end{aligned}\right\} \tag{35}$$

$$\left.\begin{aligned} \hat{\Lambda}_{0t}(t) \pm 1.96SE\left(\hat{\Lambda}_{0t}(t)\right) &\Rightarrow \sum_{j=1}^{m}\hat{\alpha}_j I_j(t) \pm 1.96\sqrt{Var\left(\sum_{j=1}^{m}\hat{\alpha}_j I_j(t)\right)} \Rightarrow I^T(t)\alpha \pm 1.96\sqrt{I^T(t)\hat{H}_{\alpha\alpha}^{-1}I(t)} \\ \hat{\Lambda}_{0c}(t) \pm 1.96SE\left(\hat{\Lambda}_{0c}(t)\right) &\Rightarrow \sum_{j=1}^{m}\hat{\eta}_j I_j(t) \pm 1.96\sqrt{Var\left(\sum_{j=1}^{m}\hat{\eta}_j I_j(t)\right)} \Rightarrow I^T(t)\eta \pm 1.96\sqrt{I^T(t)\hat{H}_{\eta\eta}^{-1}I(t)} \end{aligned}\right\} \tag{36}$$

where $I^T(t) = (I_1(t), I_2(t), \ldots, I_m(t))$ is the spline vector in t, $SE(\hat{\beta}_j)$ and $SE(\hat{\gamma}_j)$ denote the square root of the j^{th} diagonal value of the matrices $\tilde{H}_{\beta\beta}^{-1}$ and $\tilde{H}_{\gamma\gamma}^{-1}$ respectively.

3. Results and Discussions

3.1. Simulation Studies

To gain a better understanding of the empirical properties of the proposed PMLEs, Monte Carlo simulations using samples of sizes $n = 100, 500$ and 1000 with two different dependent censoring proportions $\pi_c = 20\%$ and 50% for each n were carried out. The frailty variable b_i was generated from Gamma with $\theta = 0.5$ and the failure times T_i from $\Lambda_T(t|Z_i, b_i) = b_i\Lambda_0(t_i)\,exp(\beta_1 Z_{i1} + \beta_2 Z_{i2})$ with $\Lambda_{0t}(t) = t_i$, $\beta_1 = 1$ and $\beta_2 = 0.6$. The observation time C_i^* were generated from an exponential distribution on the interval [0, 2] with an appropriately chosen exponential parameter λ to yield the desired informative censoring propor-

tions. We generated an exponential informative observation time C_i using $\Lambda_C(c_i|Z_i,b_i) = b_i\Lambda_0(c_i)\,exp(\gamma_1 Z_{i1} + \gamma_2 Z_{i2})$ with $\Lambda_{0c}(c_i) = 0.5c_i$, $\gamma_1 = 1$ and $\gamma_2 = 0.5$. Two covariates were generated from $Z_{i1}\sim Bernoulli(0.5)$ and $Z_{i2}\sim Uniform(0,1)$. Therefore, the observed data for the i^{th} individual is obtained as follows;

(i) Generate T_i, C_i and C_i^* from their distributions and define $\tilde{C}_i = min(C_i, C_i^*)$.

(ii) If $\Delta_i = 0$ and $T_i > \tilde{C}_i$, then set $\delta_i = 0 \Rightarrow (\tilde{C}_i, 0, 0, Z_{i1}, Z_{i2})$

(iii) If $\Delta_i = 1$ and $T_i > \tilde{C}_i$, then set $\delta_i = 0 \Rightarrow (\tilde{C}_i, 0, 1, Z_{i1}, Z_{i2})$

(iv) If $\Delta_i = 0$ and $T_i \le \tilde{C}_i$, then $\delta_i = 1 \Rightarrow (\tilde{C}_i, 1, 0, Z_{i1}, Z_{i2})$

(v) If $\Delta_i = 1$ and $T_i \le \tilde{C}_i$, then set $\delta_i = 1 \Rightarrow (\tilde{C}_i, 1, 1, Z_{i1}, Z_{i2})$

Cubic I-splines with knots determined based on the suggested approach was employed. The knots were placed at quantiles within $\left[\tilde{C}_{\min}, \tilde{C}_{\max}\right]$. To estimate κ_1 and κ_2, for the first replicate of each simulation setting, κ_h is estimated using the cross-validation method and this κ_h is fixed to generate the remaining 499 datasets. For each n and π_c, simulations with 500 replications were conducted. Table 1 displays the results. It can be observed that (i) for a fixed π_c, the sample standard deviations (SSD), average estimated standard errors (AESE), mean square error (MSE) and absolute value of BIAS decreases with increasing n. This phenomenon supports the assumption that the PMLEs are consistent. (ii) Comparison of values between SSD and AESE are almost equal, demonstrating that the proposed approach to estimating the variance of the PMLEs is generally accurate. (iii) the 95% coverage probabilities (CP) are reasonable since the estimates agree very well with the nominal value of 0.95 under the different $n's$ and $\pi_c's$. (iv) The mean integrated square error (MISE) of $\hat{\Lambda}_{0t}(\cdot)$ and $\hat{\Lambda}_{0c}(\cdot)$ generally shows small values, which indicates that the proposed method of estimation is generally satisfactory.

Table 1. *Simulation results on the Regression Parameters and baseline cumulative hazards of T and C for the proposed PMLE method based on 500 replications.*

n	Interior knot	Dependent censoring rate	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)	MISE
			$\beta_1 = 1$	1.161	0.161	0.606	0.557	0.393	94.6	
			$\beta_2 = 0.6$	0.716	0.116	0.915	0.865	0.851	95.3	
			$\gamma_1 = 1$	1.105	0.105	0.592	0.541	0.362	94.4	
100	4	20%	$\gamma_2 = 0.5$	0.612	0.112	0.860	0.852	0.972	93.7	
			$\theta = 0.5$	0.636	0.137	0.689	0.548	0.493	97.5	
			$\Lambda_{0t}(t)$				94.4			1.487
			$\Lambda_{0c}(t)$				95.5			0.801
			$\beta_1 = 1$	1.029	0.029	0.240	0.224	0.059	94.5	
			$\beta_2 = 0.6$	0.627	0.027	0.338	0.346	0.115	96.2	
			$\gamma_1 = 1$	0.992	- 0.008	0.221	0.224	0.049	94.6	
500	7	20%	$\gamma_2 = 0.5$	0.497	- 0.003	0.383	0.363	0.147	94.8	
			$\theta = 0.5$	0.495	- 0.005	0.274	0.214	0.075	97.6	
			$\Lambda_{0t}(t)$				94.7			1.622
			$\Lambda_{0c}(t)$				95.8			0.988
			$\beta_1 = 1$	1.021	0.021	0.161	0.158	0.026	96.0	
1000	10	20%	$\beta_2 = 0.6$	0.616	0.016	0.239	0.243	0.057	94.8	
			$\gamma_1 = 1$	1.014	0.014	0.155	0.157	0.024	96.4	

n	Interior knot	Dependent censoring rate	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)	MISE
			$\gamma_2 = 0.5$	0.507	0.007	0.257	0.254	0.066	94.8	
			$\theta = 0.5$	0.508	0.008	0.197	0.135	0.039	96.8	
			$\Lambda_{0t}(t)$				95.0			1.811
			$\Lambda_{0c}(t)$				96.8			1.058
			$\beta_1 = 1$	1.143	0.143	0.602	0.502	0.383	95.4	
			$\beta_2 = 0.6$	0.727	0.127	0.850	0.731	0.739	95.4	
			$\gamma_1 = 1$	1.043	0.043	0.405	0.392	0.165	95.8	
100	4	50%	$\gamma_2 = 0.5$	0.565	0.065	0.646	0.628	0.422	95.6	
			$\theta = 0.5$	0.584	0.084	0.514	0.404	0.272	94.4	
			$\Lambda_{0t}(t)$				95.3			1.798
			$\Lambda_{0c}(t)$				95.2			1.665
			$\beta_1 = 1$	1.036	0.036	0.227	0.210	0.053	96.6	
			$\beta_2 = 0.6$	0.606	0.006	0.308	0.302	0.095	95.0	
			$\gamma_1 = 1$	1.003	0.003	0.172	0.167	0.029	94.4	
500	7	50%	$\gamma_2 = 0.5$	0.528	0.028	0.282	0.267	0.080	94.6	
			$\theta = 0.5$	0.522	0.022	0.199	0.140	0.040	94.7	
			$\Lambda_{0t}(t)$				95.4			1.987
			$\Lambda_{0c}(t)$				93.6			1.866
			$\beta_1 = 1$	1.016	0.016	0.150	0.147	0.023	94.8	
			$\beta_2 = 0.6$	0.620	0.020	0.215	0.211	0.047	94.4	
			$\gamma_1 = 1$	1.014	0.014	0.112	0.117	0.013	96.0	
1000	10	50%	$\gamma_2 = 0.5$	0.514	0.014	0.180	0.187	0.032	95.4	
			$\theta = 0.5$	0.508	0.008	0.140	0.097	0.020	94.6	
			$\Lambda_{0t}(t)$				97.5			2.027
			$\Lambda_{0c}(t)$				94.2			2.033

Also, Figure 1 below presents a sample of histograms of parameters under the different $n's$ and $\pi_c's$ which clearly exhibit the normality distribution. In addition, it can observed in Figure 2 below that (i) the plots of the estimated and the true $\Lambda_{0t}(\cdot)$ and $\Lambda_{0c}(\cdot)$ are very close to each other under the different $n's$ and $\pi_c's$. (ii) the bounds of the 95% pointwise credible intervals for each curve contains the true curve and close to it as well, indicating the existence, if any, of small bias. Furthermore, it is observed from Tables 2-4 below that the estimated values obtained from varying number of knots chosen as 3, 5 and 8 are not significantly different from each other, revealing that the proposed method is robust to the choice of number of knots. Moreover, Table 5 displays the results of a misspecified frailty variable from the log-normal distribution with $\theta = 0.5$ and it can be noticed that the estimated parameters obtained from the misspecified distribution are very much similar to those realized from the cor-

rectly specified gamma frailty distribution. This suggest that the PMLE method proposed is robust to misspecification of the distribution of the frailty variable.

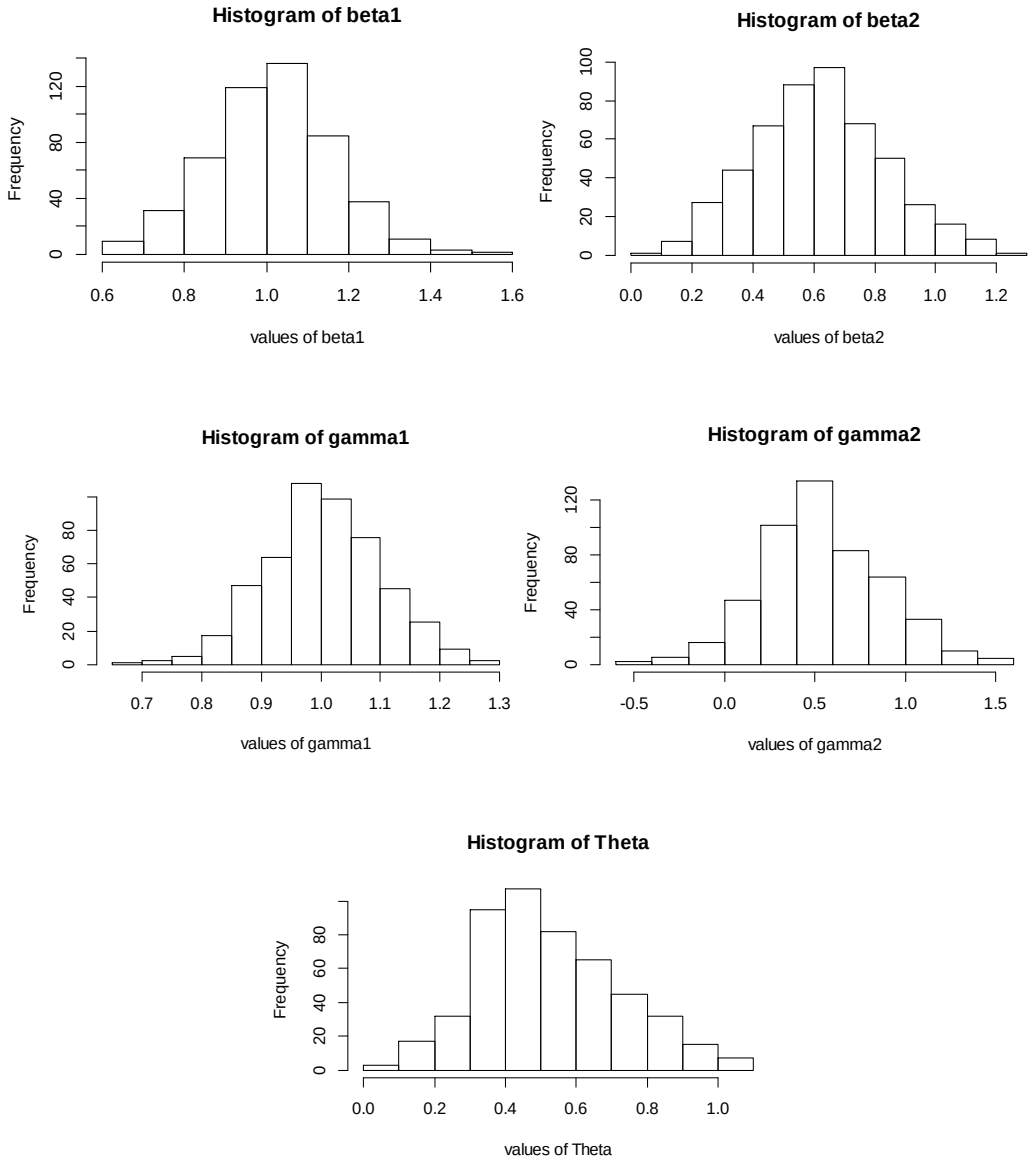

Figure 1. *Plots of Histogram for the PMLEs of the regression and frailty parameters for* $n = 500$ *with* $q_n = 7$ *and* $\pi_c = 20\%$ *based on 500 replications.*

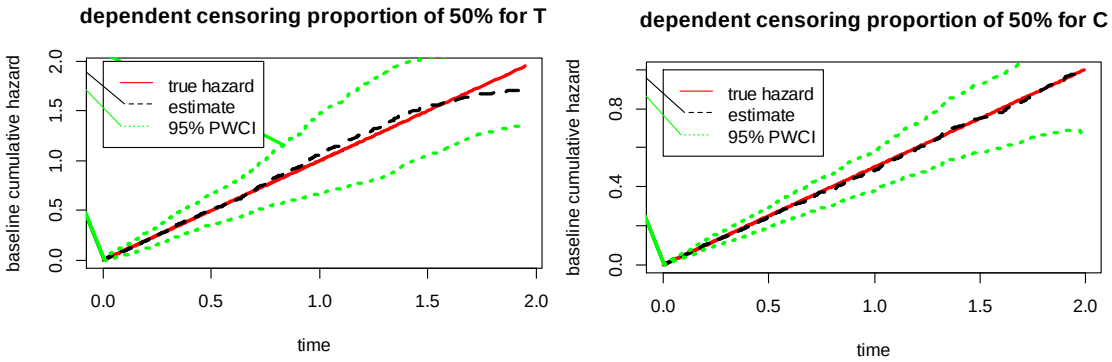

Figure 2. *Plots of true* $\Lambda_{0t}(t)$ *and* $\Lambda_{0c}(t)$ *(red solids), the PMLEs (black dashes) and 95% PWCI (green dotted), for* $n = 100$ *and* $\pi_c = 50\%$ *based on 500 replications.*

Table 2. *Simulation results on the Regression Parameters of T and C with different knots for n=100 based on 500 replications.*

Interior knot	Dependent censoring rate	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)
		$\beta_1 = 1$	1.180	0.180	0.716	0.552	0.545	94.3
		$\beta_2 = 0.6$	0.867	0.267	0.972	0.855	1.016	94.4
3	20%	$\gamma_1 = 1$	1.128	0.128	0.576	0.543	0.348	95.9
		$\gamma_2 = 0.5$	0.628	0.128	0.822	0.848	0.693	96.4
		$\theta = 0.5$	0.616	0.116	0.688	0.566	0.486	94.0
		$\beta_1 = 1$	1.183	0.183	0.584	0.556	0.374	95.6
		$\beta_2 = 0.6$	0.779	0.179	0.987	0.866	1.006	94.6
5	20%	$\gamma_1 = 1$	1.098	0.098	0.592	0.538	0.360	95.1
		$\gamma_2 = 0.5$	0.627	0.127	0.948	0.861	0.915	94.8
		$\theta = 0.5$	0.643	0.143	0.705	0.556	0.518	93.9
		$\beta_1 = 1$	1.162	0.162	0.671	0.564	0.477	95.0
		$\beta_2 = 0.6$	0.829	0.229	1.015	0.872	1.082	94.5
8	20%	$\gamma_1 = 1$	1.107	0.107	0.587	0.550	0.356	94.7
		$\gamma_2 = 0.5$	0.651	0.151	0.954	0.859	0.933	95.7
		$\theta = 0.5$	0.681	0.181	0.814	0.569	0.696	94.8
		$\beta_1 = 1$	1.150	0.150	0.527	0.497	0.300	94.7
		$\beta_2 = 0.6$	0.794	0.194	0.778	0.734	0.643	95.2
3	50%	$\gamma_1 = 1$	1.062	0.062	0.408	0.391	0.170	95.8
		$\gamma_2 = 0.5$	0.605	0.105	0.671	0.624	0.462	95.3
		$\theta = 0.5$	0.572	0.072	0.458	0.391	0.215	94.1
		$\beta_1 = 1$	1.178	0.178	0.547	0.507	0.330	95.4
		$\beta_2 = 0.6$	0.800	0.200	0.784	0.739	0.656	94.6
5	50%	$\gamma_1 = 1$	1.053	0.053	0.408	0.393	0.169	95.2
		$\gamma_2 = 0.5$	0.587	0.087	0.658	0.626	0.441	96.0
		$\theta = 0.5$	0.575	0.075	0.447	0.381	0.205	93.7
		$\beta_1 = 1$	1.227	0.227	0.569	0.528	0.375	95.4
		$\beta_2 = 0.6$	0.767	0.167	0.835	0.756	0.725	96.2
8	50%	$\gamma_1 = 1$	1.103	0.103	0.408	0.403	0.177	96.0
		$\gamma_2 = 0.5$	0.587	0.087	0.673	0.632	0.460	95.0
		$\theta = 0.5$	0.606	0.106	0.448	0.366	0.212	92.8

Table 3. *Simulation results on the Regression Parameters of T and C with different knots for* $n = 500$ *based on 500 replications.*

Interior knot	Dependent censoring rate	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)
		$\beta_1 = 1$	1.035	0.035	0.226	0.225	0.052	96.2
		$\beta_2 = 0.6$	0.623	0.023	0.346	0.348	0.120	95.6
3	20%	$\gamma_1 = 1$	1.010	0.010	0.223	0.224	0.050	95.2
		$\gamma_2 = 0.5$	0.545	0.045	0.366	0.363	0.136	94.5
		$\theta = 0.5$	0.530	0.030	0.270	0.202	0.074	93.2
		$\beta_1 = 1$	1.024	0.024	0.220	0.224	0.049	95.0
		$\beta_2 = 0.6$	0.629	0.029	0.338	0.346	0.115	95.8
5	20%	$\gamma_1 = 1$	1.010	0.010	0.216	0.223	0.047	95.4
		$\gamma_2 = 0.5$	0.511	0.011	0.348	0.362	0.121	96.6
		$\theta = 0.5$	0.527	0.027	0.255	0.199	0.066	92.8
		$\beta_1 = 1$	1.041	0.041	0.239	0.226	0.059	94.8
		$\beta_2 = 0.6$	0.622	0.022	0.349	0.349	0.122	95.2
8	20%	$\gamma_1 = 1$	1.036	0.036	0.231	0.225	0.054	94.6
		$\gamma_2 = 0.5$	0.534	0.034	0.357	0.364	0.128	95.8
		$\theta = 0.5$	0.507	0.007	0.254	0.207	0.0064	93.6
		$\beta_1 = 1$	1.029	0.029	0.205	0.209	0.043	96.4
		$\beta_2 = 0.6$	0.640	0.040	0.323	0.301	0.106	94.6
3	50%	$\gamma_1 = 1$	1.013	0.013	0.170	0.167	0.029	95.8
		$\gamma_2 = 0.5$	0.507	0.007	0.280	0.266	0.078	94.6
		$\theta = 0.5$	0.512	0.012	0.197	0.142	0.039	93.2
		$\beta_1 = 1$	1.029	0.029	0.204	0.209	0.042	94.8
		$\beta_2 = 0.6$	0.621	0.021	0.294	0.301	0.087	96.0
5	50%	$\gamma_1 = 1$	1.015	0.015	0.172	0.167	0.030	95.2
		$\gamma_2 = 0.5$	0.528	0.028	0.268	0.267	0.073	95.8
		$\theta = 0.5$	0.515	0.015	0.186	0.139	0.035	92.8
		$\beta_1 = 1$	1.006	0.006	0.157	0.157	0.025	95.0
		$\beta_2 = 0.6$	0.612	0.012	0.245	0.242	0.060	95.6
8	50%	$\gamma_1 = 1$	0.992	- 0.008	0.146	0.157	0.021	96.0
		$\gamma_2 = 0.5$	0.512	0.012	0.255	0.255	0.065	95.0
		$\theta = 0.5$	0.513	0.013	0.191	0.136	0.036	94.8

Table 4. *Simulation results on the Regression Parameters of T and C with different knots for* $n = 1000$ *based on 500 replications.*

Interior knot	Dependent censoring rate	True Parameters	AEST	BIAS	SSD	AESE	MSE	CP (%)
		$\beta_1 = 1$	1.006	0.006	0.166	0.157	0.0028	94.6
		$\beta_2 = 0.6$	0.615	0.015	0.250	0.243	0.063	94.8
3	20%	$\gamma_1 = 1$	1.016	0.016	0.163	0.158	0.027	94.5
		$\gamma_2 = 0.5$	0.496	-0.004	0.266	0.255	0.071	94.6
		$\theta = 0.5$	0.525	0.025	0.190	0.135	0.037	93.2
		$\beta_1 = 1$	1.018	0.018	0.159	0.157	0.026	94.7
		$\beta_2 = 0.6$	0.606	0.006	0.236	0.242	0.056	95.8
5	20%	$\gamma_1 = 1$	1.004	0.004	0.148	0.157	0.022	96.0
		$\gamma_2 = 0.5$	0.504	0.004	0.258	0.254	0.067	94.6
		$\theta = 0.5$	0.501	0.001	0.186	0.136	0.035	92.8
		$\beta_1 = 1$	1.006	0.006	0.157	0.157	0.025	95.0
		$\beta_2 = 0.6$	0.612	0.012	0.245	0.242	0.060	95.6
8	20%	$\gamma_1 = 1$	0.992	-0.008	0.146	0.157	0.021	96.0
		$\gamma_2 = 0.5$	0.512	0.012	0.255	0.255	0.065	95.0
		$\theta = 0.5$	0.513	0.013	0.191	0.136	0.036	94.8
		$\beta_1 = 1$	1.026	0.026	0.149	0.147	0.023	94.8
		$\beta_2 = 0.6$	0.600	0.000	0.214	0.210	0.046	94.5
3	50%	$\gamma_1 = 1$	1.008	0.008	0.114	0.117	0.013	96.2
		$\gamma_2 = 0.5$	0.517	0.017	0.192	0.187	0.037	95.0
		$\theta = 0.5$	0.507	0.007	0.134	0.097	0.018	93.6
		$\beta_1 = 1$	1.009	0.009	0.136	0.146	0.019	96.4
		$\beta_2 = 0.6$	0.600	0.000	0.211	0.210	0.045	95.2
5	50%	$\gamma_1 = 1$	1.009	0.009	0.120	0.117	0.014	95.2
		$\gamma_2 = 0.5$	0.494	-0.006	0.188	0.187	0.036	95.2
		$\theta = 0.5$	0.510	0.010	0.140	0.097	0.020	94.0
		$\beta_1 = 1$	1.011	0.011	0.151	0.147	0.023	93.6
		$\beta_2 = 0.6$	0.617	0.017	0.208	0.210	0.044	95.0
8	50%	$\gamma_1 = 1$	1.012	0.012	0.114	0.118	0.013	95.8
		$\gamma_2 = 0.5$	0.507	0.007	0.191	0.188	0.036	95.2
		$\theta = 0.5$	0.514	0.014	0.143	0.097	0.021	93.8

Table 5. *Simulation results on the frailty and regression parameters of T and C with misspecified frailty distribution based on 500 replications.*

True parameter	Statistical quantities	n=100		n=500		n=1000	
		Dependent censoring rate of 20%					
		q_n=4		q_n=7		q_n=10	
Frailty distribution		Gamma	L-Normal	Gamma	L-Normal	Gamma	L-Normal
$\beta_1 = 1$	AEST	1.161	1.194	1.029	1.026	1.019	1.025
	BIAS	0.161	0.194	0.029	0.026	0.019	0.025
	SSD	0.606	0.641	0.240	0.204	0.155	0.150
	AESE	0.557	0.542	0.224	0.218	0.157	0.152
	MSE	0.393	0.448	0.059	0.042	0.025	0.023
	CP (%)	94.0	93.0	95.4	95.1	96.8	96.0
$\beta_2 = 0.6$	AEST	0.716	0.670	0.627	0.633	0.614	0.598
	BIAS	0.116	0.070	0.027	0.033	0.014	-0.002
	SSD	0.915	0.827	0.338	0.331	0.238	0.237
	AESE	0.865	0.821	0.346	0.335	0.243	0.233
	MSE	0.851	0.690	0.115	0.110	0.057	0.056
	CP (%)	94.3	93.6	96.2	95.8	95.8	94.8
$\gamma_1 = 1$	AEST	1.105	1.140	0.992	1.033	1.012	0.999
	BIAS	0.105	0.140	-0.008	0.033	0.012	-0.001
	SSD	0.592	0.635	0.221	0.238	0.157	0.156
	AESE	0.541	0.543	0.224	0.227	0.158	0.159
	MSE	0.362	0.422	0.049	0.058	0.025	0.024
	CP (%)	94.4	95.4	94.6	94.2	95.6	94.4
$\gamma_2 = 0.5$	AEST	0.612	0.582	0.497	0.496	0.509	0.510
	BIAS	0.112	0.082	-0.003	-0.004	0.009	0.010
	SSD	98.0	0.927	0.383	0.384	0.258	0.260
	AESE	0.852	0.836	0.363	0.362	0.255	0.255
	MSE	0.972	0.866	0.147	0.147	0.067	0.068
	CP (%)	94.2	94.3	94.8	93.5	94.2	97.2
$\theta = 0.5$	AEST	0.636	0.404	0.495	0.262	0.510	0.236
	BIAS	0.137	-0.096	-0.005	-0.238	0.010	-0.264
	SSD	0.689	0.566	0.274	0.208	0.177	0.161
	AESE	0.548	0.652	0.214	0.348	0.134	0.215
	MSE	0.493	0.330	0.075	0.100	0.031	0.095
	CP (%)	92.5	91.5	97.6	97.2	94.5	97.6
Dependent censoring rate of 50%							
$\beta_1 = 1$	AEST	1.143	1.127	1.036	1.020	1.016	1.008

True parameter	Statistical quantities	n=100		n=500		n=1000	
		Dependent censoring rate of 20%					
		q_{in}=4		q_{in}=7		q_{in}=10	
Frailty distribution		**Gamma**	**L-Normal**	**Gamma**	**L-Normal**	**Gamma**	**L-Normal**
	BIAS	0.143	0.127	0.036	0.020	0.016	0.008
	SSD	0.602	0.499	0.227	0.199	0.150	0.132
	AESE	0.502	0.463	0.210	0.196	0.147	0.137
	MSE	0.383	0.265	0.053	0.040	0.023	0.017
	CP (%)	95.4	94.8	93.6	95.4	94.8	96.4
	AEST	0.727	0.702	0.606	0.632	0.620	0.619
	BIAS	0.127	0.102	0.006	0.032	0.020	0.019
	SSD	0.850	0.752	0.308	0.281	0.215	0.199
$\beta_2 = 0.6$	AESE	0.731	0.675	0.302	0.277	0.211	0.194
	MSE	0.739	0.577	0.095	0.080	0.047	0.040
	CP (%)	95.4	94.4	95.0	94.6	95.4	95.6
	AEST	1.043	1.056	1.003	0.991	1.014	0.993
	BIAS	0.043	0.056	0.003	-0.009	0.014	-0.007
	SSD	0.405	0.350	0.172	0.147	0.112	0.113
$\gamma_1 = 1$	AESE	0.392	0.352	0.167	0.154	0.117	0.108
	MSE	0.165	0.126	0.029	0.022	0.013	0.013
	CP (%)	95.8	94.3	94.4	96.0	96.0	94.8
	AEST	0.565	0.580	0.528	0.494	0.514	0.502
	BIAS	0.065	0.080	0.028	-0.006	0.014	0.002
	SSD	0.646	0.575	0.282	0.248	0.180	0.172
$\gamma_2 = 0.5$	AESE	0.628	0.549	0.267	0.244	0.187	0.172
	MSE	0.422	0.336	0.080	0.062	0.032	0.030
	CP (%)	95.6	94.4	93.6	94.0	95.4	94.4
	AEST	0.584	0.260	0.522	0.218	0.508	0.212
	BIAS	0.084	-0.240	0.022	-0.282	0.008	-0.288
	SSD	0.514	0.312	0.199	0.140	0.140	0.105
$\theta = 0.5$	AESE	0.404	0.493	0.140	0.202	0.097	0.128
	MSE	0.272	0.155	0.040	0.099	0.020	0.094
	CP (%)	93.4	93.6	92.7	92.0	93.0	92.4

3.2. Practical Illustration

The proposed PMLE procedure was applied to the tumorigenicity experiment (see [10]) that involves bladder and lung tumors of 671 female mice. The objective of the experiment is to ascertain whether exposure of mice to suspected agents accelerates the development of lung and bladder tumors in them. The mice were assigned randomly to either a control

group or High-dose group (carcinogen 2 – acetyl amino fluorine). Each animal was examined at either natural death or sacrifice times to ascertain the current state of the two tumors. Presence of the tumor(s) means that the onset time(s) is/are left censored and right censored if otherwise. At the end of the 33 months of follow-up study period, 121 mice had left censored lung tumor onset times (69 under the control group and 52 under the high-dose group). For bladder tumors data, 124 mice had left-censored onset times (13 under the control group and 111 under the high-dose group).

Define the examination time (in months) as $\tilde{C}_i = min(C_i, C_i^*)$, where C_i denotes natural death time and assumed informative and C_i^* denotes the terminal or random sacrifice time and assumed noninformative. Consider a single binary covariate Z_i (treatment group), where $Z_i = 0$ for mice assigned to the control group and $Z_i = 1$ for those assigned to the high-dose group. The analysis of the lung tumor dataset and the bladder tumor datasets were carried out separately since univariate current status data is what is being dealt with. To estimate $\Lambda_{0t}(.)$ and $\Lambda_{0c}(.)$, the cubic I – splines was used with 8 interior knots placed at quantiles of $\tilde{C}_i$. The κ_1 and κ_2 were determined using the CVS. Table 6 below presents the results. It can be observed that the PMLE of the effect of environment on growth of lung tumor gave a p-value of 0.107, revealing that the high-dose treatment has no significant effect on lung tumor development. However, for bladder tumor onset time, the p-value is 0.0001, suggesting that the high-dose treatment has a significant effect on bladder tumor onset time. For the death hazard function accompanying the lung tumors, the PMLE gives a p-value of 0.006, suggesting that death rate among mice under the high-dose group is higher than those under the control group. Similarly, the PMLE for the hazard of death accompanying the bladder tumor gave a p-value of 0.0001 indicating that death rate among mice under the high-dose group is higher than those under the control group. In addition, the frailty parameter for lung and bladder tumors onset times are all significant with p-values respectively as 0.034 and 0.0001, indicating that the time of death of mice has a stronger correlation with the appearance of bladder than that of the lung tumors.

Table 6. *Results of parameter estimates of carcinogen 2-acetylaminofluorence effect on lung and bladder tumors onset times based on a Gamma frailty.*

Statistics	**Lung tumor** **Regression and frailty parameters**		
	β	γ	θ
Estimate	0.326	0.690	0.375
SE	0.202	0.252	0.177
95% CI	(-0.070, 0.722)	(0.196, 1.184)	(0.028, 0.722)
t - value	1.612	2.737	2.120
p – value	0.107	0.006	0.034
Bladder tumor			
Estimate	3.855	0.761	0.627
SE	0.436	0.257	0.162
95% CI	(3.000, 4.710)	(0.257, 1.265)	(0.308, 0.945)
t - value	8.838	2.960	3.860
p – value	< 0.0001	0.003	0.0001

Also, the 95% credible intervals for the baseline CHFs of the lungs tumor is shown in Figure 3 below. It can be seen in the top plot that the estimated curve lies within the interval, and that the hazard of developing lung tumor is virtually nonexistent in the first 20 months. Same observation can be seen in the hazard of dying from the lung tumor curve (bottom plot). After the 20th month, the hazard of developing the lung tumor and dying from it began to increase with time.

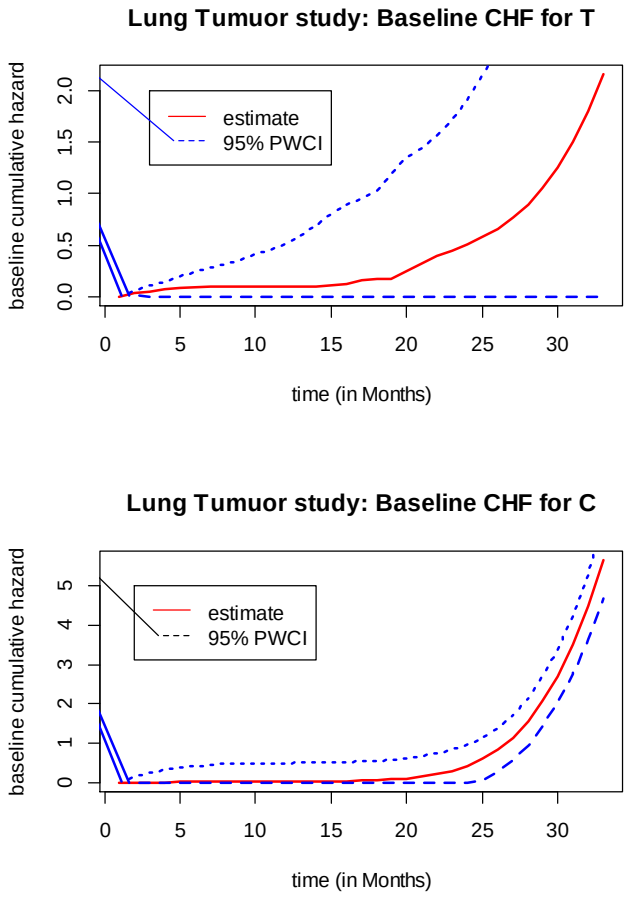

Figure 3. *Plots of estimates for lung tumor data, top is* $\Lambda_{0t}(t)$ *and bottom is* $\Lambda_{0c}(t)$ *and 95% pointwise credible interval (PWCI).*

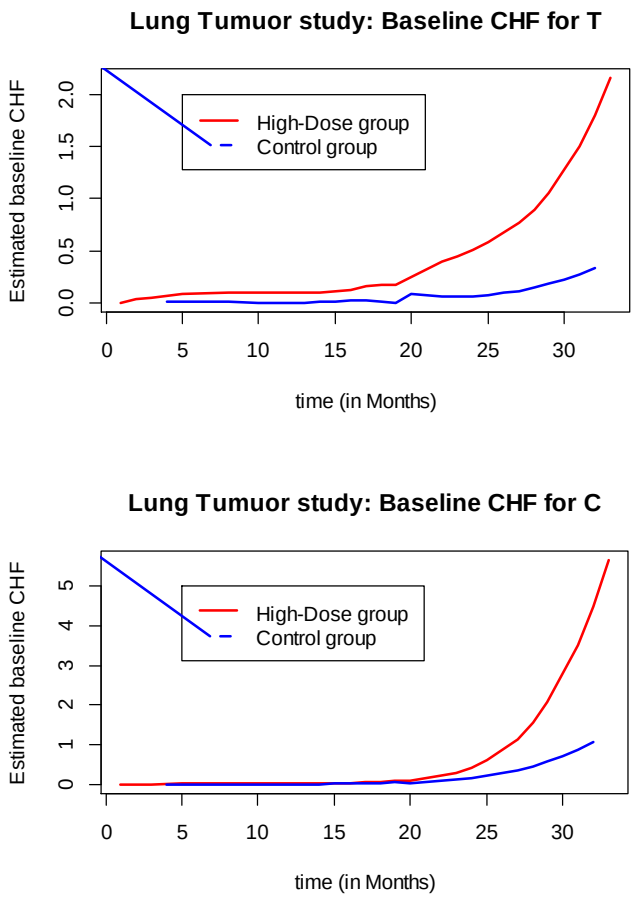

Figure 4. *Plots of baseline CHF for lung tumor onset time* $\Lambda_{0t}(t)$ *(top plot) and that of the death time* $\Lambda_{0C}(t)$ *(bottom plot), indicating treatment effect.*

Furthermore, Figure 4 below displays the baseline CHFs showing the effects of the treatment on lung tumor onset and death times. For the lung tumor onset plot (top), it can be seen that the baseline CHF for the mice in the control group is lying below that of the function for mice in the high-dose group. This means that the mice in the control group had lower hazard of developing lung tumor than those in the high-dose group. It can also be seen that the hazard of developing lung tumor for mice under the high-dose group increases with time whiles that of those under the control group virtually remains constant. For the hazard of death plot (bottom), it can be observed that the hazard of dying for mice in either the control group or high-dose group is virtually not possible within the first 20 months. However, after the 20^{th} month, whiles the hazard of dying for mice in the control group continue to exhibit constant hazard, that of those under the high-dose group began to increase sharply with time. Similar observation was seen for the bladder tumor study (though not shown here).

4. Conclusion

In this paper, a PMLE approach was proposed for analyzing informative current status data under the Cox gamma-frailty PH model, where the frailty variable was used to model the possible dependence between the failure and the observation times. To obtain the PMLEs, we proposed an easy-to-implement algorithm that utilizes BFGS Quasi-Newton algorithm. The baseline CHFs for the failure and censoring times were approximated using the cubic I-splines and smooth curves were actually realized. This made the trend of the changing hazard curves clearer for easy interpretation. The proposed method of optimal selection of knots and smoothing parameters proved to be satisfactory and therefore can be adopted by other investigators. It was also observed from the simulation studies that the PMLEs were consistent, asymptotically normal and efficient. In addition, our estimators were robust to the choice of knots, level of dependent censoring and the type of frailty distribution used. The proposed PMLE was applied to real data on tumorigenicity experiment and the findings coincide very well with previous findings (see [10]). Even though this proposed approach performed satisfactorily, some extensions can still be made in

future works. In developing the PMLE method, it was assumed the covariates to time-independent but this assumption may be violated in some survival data sets. It is however, possible to extend this proposed method to take care of time-dependent covariates.

Abbreviations

PMLE	Penalized Maximum Likelihood Estimation
CHFs	Baseline Cumulative Hazard Functions
EM	Expectation Maximization
CVS	Cross Validation Score
BFGS	Broyden Fletcher Goldfarb Shanno
CLT	Central Limit Theorem
WLLN	Weak Law of Large Numbers
SSD	Sample Standard Deviations
AESE	Average Estimated Standard Errors
MSE	Mean Square Error
CP	Coverage Probabilities
MISE	Mean Integrated Square Error

Author Contributions

Alhassan Faisal: Conceptualization, Resources, Data curation, Software, Formal Analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review & editing

Samuel Iddi: Conceptualization, Resources, Data curation, Software, Formal Analysis, Supervision, Validation, Visualization, Methodology, Project administration, Writing – review & editing

Ezekiel Nii Noye Nortey: Formal Analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review & editing

Kwabena Doku-Amponsah: Formal Analysis, Supervision, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Chen, C. M., Lu, T. F. C., Chen, M. H. and Hsu, C. M. (2012). Semiparametric transformation models for current status data with informative censoring. *Biometrical Journal*, 54, 641–656.

[2] De Boor, C. (1978). *A Practical Guide to Splines*. Springer, New York.

[3] Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, 11, 89–121.

[4] Grover, G. and Deka, B. (2013). Spline-based hazards regression model for current status data: An application to simulated data on renal impairment. *Indian Journal of Applied Research*. 3, 534 – 537.

[5] Hougaard, P. (2000). *Analysis of multivariate survival data*. Springer: New York.

[6] Klein, J. P., Moeschberger, M. L., Li, Y. H. and Wang, S. T. (1992). Estimating random effects in the Framingham heart study, Survival Analysis: State of the Art, *Kluwer Academic: Boston, Massachusetts*, 99–120.

[7] Kim, Y-J., Kim, J., Nam, C., M., and Kim, Y. (2013). Statistical analysis of dependent current status data with application to tumorigenicity experiments. *Taylor & Francis Group, LLC*.

[8] Lu, M. and Li, C. (2017). Penalized estimation for proportional hazards model with current status data. *Statistics in Medicine*, 36, 4893–4907.

[9] Li, S., Hu, T., Wang, P. and Sun, J. (2017). Regression analysis of current status data in the presence of dependent censoring with applications to tumorigenicity experiments. *Computational Statistics and Data Analysis*, 110, 75-86.

[10] Lindsey, J. C. and Ryan, L. M. (1994). A comparison of continuous- and discrete-time three-state models for rodent tumorigenicity experiments. *Environmental Health Perspectives Supplements*, 102, 9–17.

[11] Ma, L., Hu, T. and Sun, J. (2015). Sieve maximum likelihood regression analysis of dependent current status data. *Biometrika*, 102, 731–738.

[12] Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization.* Springer, New York.

[13] O'Sullivan, F. (1988). Fast computation of fully automated log-density and log-hazard estimators. *SIAM Journal of Science and Statistical Computation*. 9, 363–379.

[14] Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical science*, 3, 425–441.

[15] Vaupel, J., Manton, K. and Stallard, E. (1979). The impact of heterogeneity in individual frailty in the dynamics of mortality. *Demography*, 16, 439–454.

[16] Wahba, G. (1983). Bayesian confidence intervals for the cross-validated smoothing spline. *Journal of the Royal Statistical Society. Series B*, 45, 133–150.

[17] Wang, L., McMahan, C., Hudgens, M. and Quresh, Z. (2016). A Flexible, Computationally Efficient Method for Fitting the PH Model to Interval-Censored Data. *Biometrics*, 72, 222–231.

[18] Wang, W. and Yan, J. (2018). splines2: An R package for computing Regression Spline Functions and Classes. Version 0.2.8.

[19] Zhang, Z., Sun, J. and Sun, L. (2005). Statistical analysis of current status data with informative observation times. *Statistics in Medicine*, 24, 1399–1407.