Review Article

Journal of Future Medicine and Healthcare Innovation

Statistical Analysis of Medical Lifetime Data in Presence of Censored Data and Covariates Using Semiparametric Transformation Models Under a Bayesian Approach

Emerson Barili* and Jorge Alberto Achcar

Medical School University of São Paulo, Av. Bandeirantes, Monte Alegre, 14049-900, Ribeirão Preto, São Paulo, Brazil

***Corresponding Author**
Emerson Barili, Medical School University of São Paulo, São Paulo, Brazil.

Submitted: 2024, Feb 09; **Accepted:** 2024, Mar 04; **Published:** 2024, Mar 12

Citation: Barili, E., Achcar, J. A. (2024). Statistical Analysis of Medical Lifetime Data in Presence of Censored Data and Covariates Using Semiparametric Transformation Models Under a Bayesian Approach. *J Future Med Healthcare Innovation, 2*(1), 01-15.

Abstract
A very promising alternative recently considered in the literature for the analysis of lifetime data in presence of covariates and censored data is given by the class of semiparametric or transformation models. This class of models generalizes the usual proportional hazards models, the proportional odds models, and the accelerated failure time models, extensively used in lifetime data analysis. In the analysis of lifetime data, especially in medicine, the proportional hazards model has been the most used model due to its flexibility without the need to assume a parametric model for the data [1]. Despite this advantage, in some applications the needed assumption (proportional hazards) may not be verified and the class of transformation models can be quite attractive in data analysis. In obtaining inferences of interest, especially obtaining point estimators for the regression parameters assuming transformation models, several proposals have been introduced in the literature, as alternatives to the use of the partial likelihood proposed assuming proportional hazards models [1, 2]. In this work, we introduce a simple method to obtain inferences for the regression parameters of semiparametric models or transformation models under a Bayesian approach considering the unknown hazard rates as latent variables. The posterior summaries of interest are obtained using existing MCMC (Markov Chain Monte Carlo) simulation methods. An application with real-time medical data illustrates the proposed methodology.

Keywords: Lifetime Data, Semiparametric Models, Censored Data, Covariates, Bayesian Analysis, MCMC Methods

1. Introduction

In many applications, especially in medicine or engineering studies, we have lifetime data in presence of censored observations and covariates associated with each individual or unity. In the analysis of lifetime data, different parametric or non-parametric regression models were proposed in the literature to analyse the data, usually in the presence of censoring and covariates [3-5]. In this way, three classes of models have been extensively used in lifetime data analysis in the presence of censoring and covariates: the proportional hazards or PH models, the proportional odds or PO models, and accelerated failure time or AFT models [1, 6, 7]. In many situations, the needed assumptions for each model could be not verified, especially assuming the PH model when there are crossing survival curves (usually Kaplan and Meier estimates) assuming categorized covariates, that is, the assumption of the Cox PH model is not verified. To circumvent the lack of proportional hazards, the literature presents several studies, considering generalizations of these models, including the PH and PO models. One of these generalizations, is given by the semiparametric two-sample strategy (YP model) proposed by [8]. A unified approach to fit the YP model is introduced using Bernstein polynomials to manage the baseline hazard and odds under both the frequentist and Bayesian frameworks [9].

A class of models also widely used in the analysis of survival data is given by the class of accelerated failure time models where the effects of covariates are assumed in a linear form, which can be restrictive for many practical problems [3]. Considering more flexible nonlinear structures to model relationships between covariates and transformed failure times, proposes a class of semiparametric models in the analysis of lifetime data [10].

The proportional hazards (PH) model introduced is a semi-parametric model of simple interpretation where the occurrence of censorship is easily accommodated [1]. In addition, it is available in most statistical software. The Cox model assumes that the hazard function can be written in the form,

$$h(t; z) = h_0(t)e^{\beta z} \tag{1}$$

where t denotes the lifetime (a value of the random variable T > 0) of an individual, $h_0(t)$ is a non-negative arbitrary baseline hazard function defined by,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}, \tag{2}$$

where β is a vector of regression coefficients and z is a vector of covariates. In (1), the covariates affect the hazard function in a multiplicative way according to the factor $e^{\beta z}$. A special likelihood function was proposed (denote as a partial likelihood) that does not depend on the baseline hazard function $h_0(t)$, thus allowing inferences on β not needing to specify a parametric form for the hazard function $h_0(t)$ [1]. Under fairly weak regularity conditions, the usual asymptotic properties of likelihood-based inference are verified [2].

Many other generalizations of the proportional hazards' models were introduced in the literature in recent years. A model averaging method to produce model-based prediction for survival outcomes defined as a semiparametric model averaging prediction (SMAP) method which approximates the underlying unstructured nonparametric regression function by a weighted sum of low-dimensional nonparametric sub models was introduced by [11]. The weights are obtained from maximizing the partial likelihood constructed for the aggregated model and theoretical properties are discussed for the estimated model weights. A semiparametric survival analysis via Dirichlet process mixtures of the First Hitting Time (FHT) model was introduced, considering several random effects specifications of the FHT model under a Bayesian approach [12]. Semi-parametric models for longitudinal data and semiparametric transformation models for interval-censored data were also introduced in the literature [13, 14]. Introduced a study for the comparison of parametric and semiparametric survival regression models with kernel estimation considering two types of kernel smoothing and some bandwidth selection techniques [15]. An overview of semiparametric models commonly used in survival analysis, including proportional hazards model, proportional odds models and linear transformation models was introduced by [16].

Other studies related to the modelling of lifetime data in presence of covariates and censored data were introduced in the literature: [17] an additive risk model specifying that the hazard function associated with a set of possibly time-varying covariates is the sum of the baseline hazard function and the regression function of covariates in contrast to the usual proportional hazards model was introduced by [18, 19]; considered in place of the usual Cox-type intensity function for counting process commonly used to analyse recurrent event data, a time-transformed Poisson process assuming that the covariates have multiplicative effects on the mean and rate functions of the counting process; [20] considered partly linear transformation models (semiparametric regression models) applied to current status data where the unknown quantities are the transformation function, a linear regression parameter and a nonparametric regression effect showing flexible alternatives to the Cox model for current status data analysis; [21] introduced a covariate analysis of current status data showing that the method is applicable when the logit of the conditional probability of survival given the covariates is some increasing function of time plus a linear combination of the covariates; [22] investigated joint models for a time-to-event (e.g., survival) and a longitudinal response where the longitudinal data are assumed to follow a mixed effects model and a proportional hazards model depending on the longitudinal random effects and other covariates are assumed for the survival endpoint proposing a likelihood based approach that requires only the assumption that the random effects have a smooth density; [23] studied joint modelling of survival and longitudinal data where there are two regression models of interest, the first one for survival outcomes, which are assumed to follow a time-varying coefficient proportional hazards model and the second one is for longitudinal data, which are assumed to follow a random effects model proposing a local corrected score estimator and a local conditional score estimator to deal with covariate measurement error; [24] proposed a semiparametric additive rate model for modelling recurrent events in the presence of a terminal event where a general transformation model is used to model the terminal event; [25] introduced a new class of transformed hazard rate models that contains both the multiplicative hazards model and the additive hazards model as special cases; [26] proposed a general class of semiparametric transformation models with random effects to formulate the effects of possibly time-dependent covariates on clustered or correlated failure times encompassing all commonly used transformation models, including proportional hazards and proportional odds models; [27] proposed a large class of semiparametric transformation models with random effects for the joint analysis of recurrent events and a terminal event where the transformation models include proportional hazards/intensity and proportional odds models.

In this study, we introduce a Bayesian analysis of the semiparametric or transformed models assuming the complete likelihood function obtained from the transformation model considering the unknown hazard function as a latent unknown variable under a Bayesian approach.

The main goals of this study are:
(i) The introduction of a Bayesian approach for semiparametric or transformation models assuming the unknown hazard function as a latent random variable with a specified probability density function.
(ii) The elicitation of prior distributions for the regression parameters using empirical Bayesian methods.
(iii) The use of existing MCMC (Markov Chain Monte Carlo) methods to get the posterior summaries of interest.
(iv) The use of some Bayesian criteria, in special, the posterior Bayes factor to decide by the best special class of transformation model to be assumed in a lifetime data analysis in presence of covariates and censored data.

The paper is prepared as follows: section 2 introduces the class of semiparametric or transformation models; section 3 introduces the likelihood function under special classes of the semi parametric model; section 6 presents a Bayesian analysis considering the unknown hazard function as a random factor with a specified probability density function; section 8 presents some applications with real medical data sets; section 9 presents a simulation study to check the robustness of the proposed methodology to different proportions of censored data; finally, section 10 presents some concluding remarks.

2. Transformation Model

Let T denote the failure time, and let $z(\cdot)$ denote a d-vector of covariates associated to each individual. Under the semiparametric transformation model, the cumulative hazard function for T conditional on z is given by,

$$\Lambda(t; z) = G\left\{\int_0^t e^{\beta z} h(u)du\right\} \qquad (3)$$

where $G(\cdot)$ is a specific transformation function that is strictly increasing, β is a regression parameter and $\Lambda(\cdot)$ is an unknown increasing function defined $\Lambda(t) = \int_0^t h(u)du$ denoting the usual cumulative hazard function not considering the presence of the covariate vector $\mathbf{z}$ [28].

$$\Lambda(t; \mathbf{z}) = G\left\{e^{\beta z}\Lambda_0(t)\right\} \qquad (4)$$

where $\Lambda_0(t)$ is the baseline cumulative hazard function.

The class of semiparametric models has been recently used as an alternative in the analysis of lifetime data in the presence of censoring and covariates. A generalization of this class of models assuming the presence of a fraction of individuals not experiencing the event of interest (cured or non-susceptible individuals) was introduced by [29]. Other generalizations of the semiparametric model (or transformation models) were also proposed in the literature [30, 31]. A generalization of the semiparametric models to interval-censored data was introduced by and a maximum likelihood estimation approach for semiparametric transformation models in the presence of interval censored data was introduced by [28, 32]. Presented a hierarchical Bayesian approach for semiparametric models (or transformation models) assuming the unknown hazards as latent factors for semiparametric models; introduced a hierarchical Bayesian approach for semiparametric models (or transformation models) in presence of cure fraction [33, 34].

Some special cases of the semiparametric model (4) are given by

(i) If $G(x) = x$, $\Lambda(t; \mathbf{z}) = e^{\beta \mathbf{z}}\Lambda_0(t)$, where $\Lambda_0(t) = \int_0^t h_0(u)du$ (h_0 is unknown), that is, we have the proportional hazards model since $h(t; \mathbf{z}) = e^{\beta \mathbf{z}}h_0(t)$. In this case, two individuals denoted by i and j with covariates z_i and z_j have a hazard ratio given by $h(t; \mathbf{z}_i)/h(t; \mathbf{z}_j) = e^{\beta \mathbf{z}_i}h_0(t)/e^{\beta \mathbf{z}_j}h_0(t) = e^{\beta z_i}/e^{\beta z_j}$ (does not depend on t, that is, we have a proportional hazards model).

(ii) If $G(x) = \log(1+x)$, we have $\Lambda(t; \mathbf{z}) = \log\left\{1 + e^{\beta \mathbf{z}}\Lambda_0(t)\right\}$, $S(t; \mathbf{z}) = \exp\left(-\Lambda(t; \mathbf{z})\right) = \exp\left(-log\left[1 + e^{\beta \mathbf{z}}\Lambda_0(t)\right]\right) = 1/\left[1 + e^{\beta \mathbf{z}}\Lambda_0(t)\right]$ and $1 - S(t; \mathbf{z}) = e^{\beta \mathbf{z}}\Lambda_0(t)/\left[1 + e^{\beta \mathbf{z}}\Lambda_0(t)\right]$, ($S(t) = P(T > t)$ is the survival function) leading to the proportional odds ratio model, since

$$\begin{aligned} OR_i/OR_j &= \left\{S(t; \mathbf{z}_i)/\left[1 - S(t; \mathbf{z}_i)\right]\right\}/\left\{S(t; \mathbf{z}_j)/\left[1 - S(t; \mathbf{z}_j)\right]\right\} \\ &= \left\{S(t; \mathbf{z}_i)/S(t; \mathbf{z}_j)\right\}\left\{\left[1 - S(t; \mathbf{z}_j)\right]/\left[1 - S(t; \mathbf{z}_i)\right]\right\} \\ &= \left\{\left[1 + e^{\beta \mathbf{z}_j}\Lambda_0(t)\right]/\left[1 + e^{\beta \mathbf{z_i}}\Lambda_0(t)\right]\right\} \cdot \\ &\quad \left\{e^{\beta \mathbf{z}_j}\Lambda_0(t)/\left[1 + e^{\beta \mathbf{z}_j}\Lambda_0(t)\right]\right\} \cdot \\ &\quad \left\{\left[1 + e^{\beta \mathbf{z_i}}\Lambda_0(t)\right]/e^{\beta \mathbf{z_i}}\Lambda_0(t)\right\}. \end{aligned}$$

That is, $OR_i/OR_j = \left\{e^{\beta \mathbf{z}_j}\Lambda_0(t)/e^{\beta \mathbf{z}_i}\Lambda_0(t)\right\} = e^{\beta \mathbf{z}_j}/e^{\beta \mathbf{z_i}}$ (a proportional odds model).

(iii) If $G(x) = \log(1+rx)/r$ ($r \geq 0$), we have the logarithmic transformation family, with $G(x) = x$ if $r = 0$ and $G(x) = \log(1+x)$ if $r = 1$ [28]. In this case we have $\Lambda(t; \mathbf{z}_i) = \log\left\{1 + re^{\beta \mathbf{z}}\Lambda_0(t)\right\}/r$ and $S(t, \mathbf{z}) = \exp(-\Lambda(t, \mathbf{z})) = \exp\left(-\dfrac{\log(1 + re^{\beta \mathbf{z_i}}\Lambda_0(t))}{r}\right)$.

Remark 1 [35]: $\log(1+x) \approx x - x^2/2 + x^3/3 - \cdots (\mid x \mid \leq 1$ and $x \neq -1)$. In this way, $\Lambda(t; \mathbf{z}) = \log\{1 + e^{\beta \mathbf{z}}\Lambda_0(t)\} \approx e^{\beta \mathbf{z}}\Lambda_0(t)$ (the PH model) and $\Lambda(t; \mathbf{z}) = \log\{1 + re^{\beta \mathbf{z}}\Lambda_0(t)\}/r \approx e^{\beta \mathbf{z}}\Lambda_0(t)$ (the PH model).

3. Likelihood Function

The likelihood function in the presence of right-censored data and a vector of covariates ***Z*** is given by,

$$L(\cdot) = \prod_{i=1}^{n} f(t_i; z_i)^{\delta_i} S(t_i; z_i)^{1-\delta_i} \tag{5}$$

where z_i is a vector of covariates; $\delta_i = 1$ (complete observation) and $\delta_i = 0$ (censored observation). Since *h(t) = f(t)/S(t),* we have *f(t) = h(t)S(t).* Thus, the likelihood function 5 can be written as,

$$L(\cdot) = \prod_{i=1}^{n} h(t; z_i)^{\delta_i} S(t_i; z_i) \tag{6}$$

Remark 2:

(i) The accumulated hazard function Λ(t) can be given by

$$\Lambda(t) = \int_0^t h(u)du = -\log[S(t)].$$

(ii) The hazard function *h(t)* can be given by *h(t) = dΛ(t)/dt = −S'(t)/S(t).*

(iii) The density function *f(t)* can be given by *f(t) = dF(t)/dt* where *F(t)* = 1 − *S(t).*

4. The Transformation G(x) = x (A Proportional Hazards Model)

In this case, we have $\Lambda(t;\mathbf{z}) = e^{\beta\mathbf{z}}\Lambda_0(t)$ that is, $h(t;\mathbf{z}) = e^{\beta\mathbf{z}}h_0(t)$ (hazard function). Thus, $S(t;\mathbf{z}) = e{-}\Lambda(t;\mathbf{z}) = e{-}e^{\beta\mathbf{z}}\Lambda_0(t)$, and $f(t;\mathbf{z}) = h(t;\mathbf{z})$ $S(t;\mathbf{z}) = e^{\beta\mathbf{z}}h_0(t)e{-}e^{\beta\mathbf{z}}\Lambda_0(t)$ Thus, the likelihood function (in the presence of right-censored data and covariates) is given by,

$$\begin{aligned} L(\cdot) &= \prod_{i=1}^{n} h(t_i, \mathbf{z}_i)^{\delta_i} S(t_i; \mathbf{z}_i) \\ &= \prod_{i=1}^{n} \left[e^{\beta\mathbf{z}_i} h_0(t_i)\right]^{\delta_i} e^{-e^{\beta\mathbf{z}_i}\Lambda_0(t_i)} \\ &= e^{\sum\limits_{i=1}^{n}\delta_i\beta\mathbf{z}_i} \left\{\prod_{i=1}^{n} [h_0(t_i)]^{\delta_i}\right\} e^{-\sum\limits_{i=1}^{n} e^{\beta\mathbf{z}_i}\Lambda_0(t_i)}. \end{aligned} \tag{7}$$

The log-likelihood function is given by,

$$\begin{aligned} l(\cdot) &= \log[L(\cdot)] \\ &= \sum_{i=1}^{n} \delta_i\beta\mathbf{z}_i + \sum_{i=1}^{n} \delta_i \log[h_0(t_i)] - \sum_{i=1}^{n} e^{\beta\mathbf{z}_i}\Lambda_0(t_i). \end{aligned}$$

Considering the contribution of only one observation *i*, we have:

$$l(\cdot) = \log[L(\cdot)] = \delta_i\beta\mathbf{z}_i + \delta_i \log[h_0(t_i)] - e^{\beta\mathbf{z}_i[\Lambda_0(t_i)]} \tag{8}$$

where $h_0(t_i)$ and $\Lambda_0(t_i)$ are unknown.

5. The Logarithmic Transformation Family *G*(*x*) = log(1 + *rx*)/*r*

In this case we have,

$$\begin{aligned} \Lambda(t; z_i) = \log\{1 + re^{\beta\mathbf{z}}\Lambda_0(t)\}/r \;\; &\text{and} \;\; S(t; \mathbf{z}) \\ &= \exp(-\Lambda_0(t; \mathbf{z})) = \exp\left(-\frac{\log(1 + re^{\beta z_i}\Lambda_0(t))}{r}\right) \end{aligned}$$

That is,

$$S(t;\mathbf{z}) = 1/[1 + re^{\beta\mathbf{z}}\Lambda_0(t)]^{1/r} \tag{9}$$

where $\Lambda_0(t) = \int_0^t h_0(u)du$ and the probability density function $f(t;\mathbf{z}) = -dS(t;\mathbf{z})/dt$ is given by,

$$f(t;\mathbf{z}) = e^{\beta\mathbf{z}}h_0(t)/[1 + re^{\beta\mathbf{z}}\Lambda_0(t)]^{1/r+1} \tag{10}$$

and $h_0(t;\mathbf{z}) = f(t;\mathbf{z})/S(t;\mathbf{z}) = e^{\beta\mathbf{z}}h_0(t)/[1 + re^{\beta\mathbf{z}}\Lambda_0(t)]$.

From 6 the likelihood function based on the i^{th} observation is given by,

$$L(r,\beta) = \left\{e^{\beta\mathbf{z}_i}h_0(t)/[1 + re^{\beta\mathbf{z}_i}\Lambda_0(t_i)]\right\}^{\delta_i} \left\{1/[1 + re^{\beta\mathbf{z}_i}\Lambda_0(t_i)]^{1/r}\right\} \tag{11}$$

The log-likelihood function is given from (11) by,

$$l(r,\beta) = \beta\mathbf{z}_i\delta_i + \delta_i\log[h_0(t_i)] - \delta_i\log[1 + re^{\beta\mathbf{z}_i}\Lambda_0(t_i)] - \left\{\log[1 + re^{\beta\mathbf{z}_i}\Lambda_0(t_i)]\right\}/r \tag{12}$$

A special case of (12) is obtained when r =1 (proportional odds model).

6. A Bayesian Analysis Considering the Unknown Hazard Function as A Random Factor

We obtain inferences for the transformation model introduced in section (2) under a Bayesian approach [36]. Since the baseline hazard function $h_0(t)$ is unknown, we assume $h_0(t)$ as a latent unknown random variable. In this way, we assume $d_i = h_0(t_i)$ as a random effect with a gamma probability distribution $G(a,b)$ with mean a/b and variance a/b^2. Thus, the cumulative hazard function is given by, $\Lambda_0(t_i) = d_i t_i$. We use standard MCMC (Markov Chain Monte Carlo) methods as the Gibbs sampling algorithm or the Metropolis-Hastings algorithm to get the posterior summaries of interest [37, 38]. Assuming only one covariate, we also assume a gamma prior distribution for the parameter $\theta = \exp(\beta)$ that is, $\theta \sim G(c,d)$ where c and d are known hyperparameters. We assume the reparameterization $\theta = \exp(\beta)$ to have better convergence of the Gibbs sampling algorithm. Considering a vector of covariates associated with a vector of parameters $\theta = (\theta_1,\theta_2,\theta_3,\cdots,\theta_k),\theta_j = \exp(\beta_j)$,**j = 1,2,...,k**, we assume independent gamma prior distributions $\theta_j \sim G(c_j,d_j)$. Under a Bayesian approach, we use the Bayes formula to combine a specified prior distribution for the parameters of the model with the likelihood function of the model, obtaining the posterior distribution from where the Bayesian inferences are obtained. Therefore, for θ, the vector of parameters of a model describing the behaviour of the data D, if $P(\theta)$,$P(\theta \mid D)$, and $L(D \mid \theta)$ indicate, respectively, the prior, the posterior distributions of θ, and the likelihood function of the model, then $P(\theta \mid D) \propto L(D \mid \theta)P(\theta)$.

For the discrimination of the best model, we use the posterior Bayes factor [38].

It is important to point out that other different approaches were introduced in the literature for the unknown baseline hazard function $h_0(t)$ as alternative for the proposed method considered in this study. In this way, assumed to approximate the baseline hazard function by a Taylor series considering interval-censored time-to-event data, but the elicitation of appropriate prior distributions for the regression parameters under this model approach is not an easy task, where we usually have convergence problems for the iterative MCMC simulation method [29].

7. Bayesian Discrimination of the Best Model

Some discrimination criteria such as the Bayesian information criterion or BIC and the deviance information criterion or DIC could be alternatives in the selection of different models, but these standard criteria will always select the models with more parameters [39-41]. In this study, in order to select the best model, we consider a Bayesian discrimination method, given by the posterior Bayes factor where the generated Gibbs samples for the parameters of each model are used to obtain Monte Carlo estimates of the Bayes factor for the special cases of the semi parametric or transformation model [42].

The posterior Bayes factor is as a discrimination criterion between two models i and j given by $B_{ij} = V_i/V_j$ where V_k is the posterior mean of the likelihood function under model k given by, Z

$$V_k = \int L(\mathbf{D} \mid \theta_k)P(\theta_k \mid \mathbf{D})d\theta_k \tag{13}$$

$L(\mathbf{D} \mid \theta_k)$ is the likelihood function under Model k and $P(\theta_k \mid D)$ is the joint posterior distribution of the vector of parameter θ_k. Observe that V_k is the expected value of the likelihood function with respect to the joint posterior distribution for θ_k.

If $B_{ij} = V_i/V_j > 1$, then the Bayes factor criterion favors model i. Observe that the corresponding value of Vi for the ith assumed model is given by the product of the likelihood functions (Monte Carlo estimate from the Gibbs samples) considering each one of the observations in the sample. This Monte Carlo estimate is obtained directly from the OpenBUGS software [43].

8. Applications

8.1. Example 1

The remission times of 42 patients with acute leukaemia were reported in a clinical trial conducted to assess the ability of 6-mercaptopurine (6-MP) to maintain remission [44]. Each patient was randomized to receive 6-MP or a placebo. The study finished after one year. The following remission times, in weeks, were recorded: 6-MP (21 patients): 6, 6, 6, 7, 10, 13, 16, 22, 23, 6+, 9+, 10+, 11+, 17+, 19+, 20+, 25+, 32+, 32+, 34+, 35+; Placebo (21 patients): 1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23.

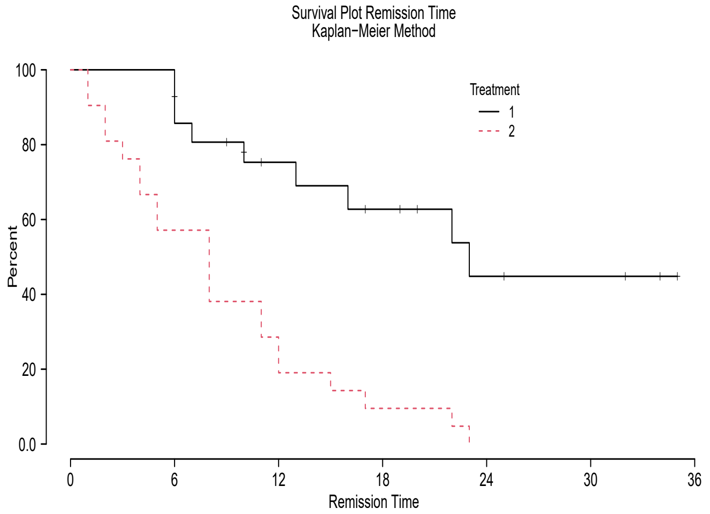

Figure 1: Kaplan-Meier Estimated Curves for the Two Groups (Example 1)

From Figure 1, we observe that the estimated non-parametric survival plots are not crossing assuming the two treatment groups, an indication that the PH model could be used in the data analysis (a subjective decision commonly assumed by medical researchers in applications to use PH models) [45]. Assuming a PH (proportional hazards) model (1) in presence of only a covariate $z(z = 1$ for the 6-MP group and $z = 2$ for the placebo group), the MLE (maximum likelihood estimator) for the regression parameter β (use of the partial likelihood function) obtained using the R software is given by $\beta_b = 1.509(0.470)$ where the value in parentheses is the standard error. That is, $\theta = \exp(1.509) = 4.52221$.

For a Bayesian analysis of the data considering special classes of the semiparametric models introduced in section 2, that is, the PH or proportional hazards model denoted as "model", the logarithmic transformation model denoted as "model 2" and the PO or proportional odds ratio model ($r = 1$) denoted as "model 3" in the logarithmic transformation model, we assume that the hazard function $h_0(t_i), i = 1,2,\cdots,n$ is an unknown latent factor, that is, $h_0(t_i) = d_i$ with a gamma distribution $G(0.1,0.1)$.

As a first model, we assume the PH model (model 1) considering a gamma $G(a,b)$ prior distribution for the parameter θ (reparameterization of β) with hyperparameters $a = 20.25$ and $b = 4.5$ where this prior distribution was elicited from prior knowledge obtained assuming the PH model under the classical approach using the partial likelihood proposed, by solving the equations, $E(\theta) = a/b = 4.5$ and $var(\theta) = a/b^2 = 1$ (mean and variance of the Gamma distribution using the prior information obtained from the obtained point estimate for $\theta = exp(\beta)$ using the PH model) [1]. Observe that $\Lambda_0(t_i) = d_i t_i$, where $h_0(t_i) = d\Lambda_0(t_i)/$ dti. Also observe that we are using empirical Bayesian methods since the prior information from the classical inferences based on the partial likelihood for the PH Cox model was used in the elicitation of prior for the regression parameter β [40].

From the approximation $\log(1 + x) \approx x$ (remark 1), we also assume the same gamma $G(20.25,4.5)$ prior distribution for θ assuming "model 2" (logarithmic transformation model) and "model 3" (PO model or proportional odds ratio model).

For all cases, we used the OpenBUGS software considering a burn-in sample of 1,000 simulated samples discarded to eliminate the effects of the initial values in the iterative procedure and taking a final sample of size 1, 000 (every 100th in 100,000 generated Gibbs samples) to get the Monte Carlo Carlo estimates for the parameters of interest [43]. The convergence of the Gibbs sampling algorithms was verified from trace plots of the simulated samples for each parameter. The OpenBUGS codes used in this application are presented in an appendix at the end of the manuscript. Table 1 shows the posterior summaries for each assumed model.

	β		θ		r		
	Mean	SD.	Mean	SD.	Mean	SD.	Bayes Factor
Model 1	1.448	0.235	4.370	1.012	—	—	$V_1 = \exp(-106.6)$
Model 2	1.473	0.226	4.474	0.984	0.376	0.662	$V_2 = \exp(-110.9)$
Model 3	1.459	0.231	4.414	1.005	—	—	$V_3 = \exp(-119.4)$

Table 1: Posterior Summaries Assuming the Transformation Models (Example 1)

From the results of Table 1, we observe that "model 1" (the PH model) is the best fitted model for the data using the posterior Bayes factor as a discrimination criterion, since the Bayesian Monte Carlo estimates for V_k (k indexes each assumed model, that is, k = 1 for "model 1", k = 2 for "model 2" and k = 3 for "model 3") in (13) are given, respectively by $V_1 = \exp(-106.6)$, $V_2 = \exp(-110.9)$ and $V_3 = \exp(-119.4)$, with larger value for V_1, an indication that "model 1" (PH model) is the best model fitted by the data set. This result is in agreement with the plots of the Kaplan-Meier curves in Figure 1 (not crossing survival curves, an indication that the PH model is adequate). Observe that the MLE estimate obtained using the PH model under the partial likelihood was given by β_b = 1.509(0.470), that is, the standard error (0.470) obtained from asymptotic results based on the partial likelihood proposed is larger than obtained using the Bayesian approach assuming "model 1", as a special case of the general semiparametric model [1]. The 95% credible interval for the parameter r (logarithmic transformation model) obtained from the simulated Gibbs samples is given by (0.00005;2.466).

8.2. Example 2

In this example, we consider the data (survival times in day) from 35 cancer patient treated at the Mayo Clinic: data from sample 1 (large tumour): 28, 69, 175, 309, 377+, 393+, 421+, 447+, 462+, 709+, 744+, 770+, 1106+, 1206+; data from sample 2 (small tumour): 34, 88, 137, 199, 280, 291, 299+, 300+, 309, 351, 358, 369, 369, 370, 375, 382, 392, 429+, 451, 1119+ [46]. Define the covariate Z denoting both sample groups: $Z = 1$ for large tumour, $Z = 0$ for small tumour. Assuming the PH model (1) , the hazard function is given by $h(t \mid x) = h_0(t)e^{\beta z}$. In this way when z equals 1 (large tumour) and z is equals 0 (small tumour), we have: $h(t \mid z = 1) = h_0(t)e^{\beta}$ if $z = 1$, and $h(t \mid z = 0) = h_0(t)$ if $z = 0$. The Cox proportional hazards model was fitted for tumour size (large and small tumours). The 95% confidence interval for e^{β} is given by (0.123; 0.865). From the obtained inference results, we observe a statistically significant difference in survival times between the two groups. The large tumour group has the highest survival times. The MLE estimator for e^{β} using the partial likelihood function is given by 0.327 and the 95% confidence interval is given by (0.123; 0.865). The MLE estimate of the β regression parameter is given by -1.119. This means that the probability of surviving the event is higher in the large tumour group, that is, the risk of failure is 67.3% lower in the large tumour group when compared to the small tumour group.

From Figure 2, showing the Kaplan-Meier estimates for the survival functions in the two groups, we observe that in this example it is hard to decide that the PH model is appropriate for the data analysis based on the estimated Kaplan-Meier curves.

A reanalysis of the data is considered assuming the semiparametric models introduced in section 2.

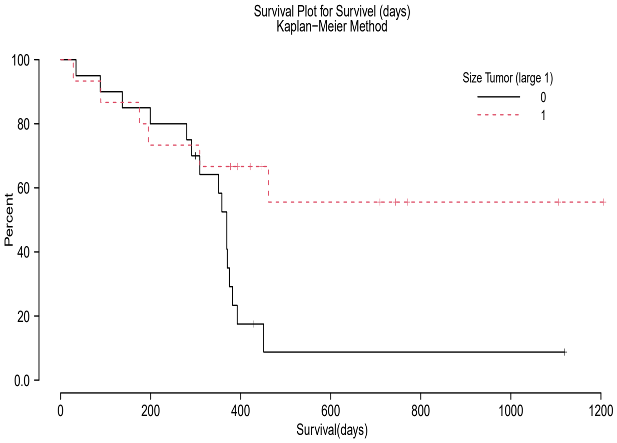

Figure 2: Kaplan-Meier Estimated Curves for the Two Groups (Example 2)

For a Bayesian analysis of the data considering the semiparametric models introduced in section 2 (the proportional hazards model denoted as "model 1", the logarithmic transformation model denoted as "model 2" and the proportional odds ratio model (r = 1) denoted as "model 3" in the logarithmic transformation model) we also assume that the hazard function $h_0(t_i), i = 1,2,\ldots,n$ is an unknown latent factor, with a gamma distribution $G(0.1,0.1)$. For "model 1" (PH model), we assume a gamma $G(1.07,3.27)$ prior distribution for the parameter θ where the elicitation of the hyperparameters in the gamma prior distribution for the parameter was obtained by solving the equations $E(\theta) = a/b = 0.327$ and $\text{var}(\theta) = a/b^2 = 0.1$. We used the prior information of the Cox PH model (β estimated by -1.119), that is, $\exp(-1.119) = 0.327$. Assuming "model 2" (logarithmic transformation model) and "model 3" (PO model), we assumed the same gamma prior distribution $G(1.07,3.27)$ for the parameter θ and a non-informative gamma $G(0.1,0.1)$ prior distribution for the parameter r in "model 2" (logarithmic transformation model). We used the same Gibbs simulation procedure considered in example 1 to simulate samples of the joint posterior distribution of interest. Table 2 shows the posterior summaries for each assumed model.

From the results of Table 2, we have the Bayesian Monte Carlo estimates (13), $V_1 = exp(-153.3)$, $V_2 = exp(-156.3)$ and $V_3 = \exp(-162.8)$, an indication that "model 1" (PH model) possibly is the best model fitted by the data set (larger value for V_1). A 95% credible interval for the parameter r (logarithmic transformation model or "model 2") obtained from the simulated Gibbs samples is given by (1.3; 1.38). Although the posterior Bayesian criterion indicates "model 1", since the Monte Carlo estimates for V_1 and V_2 are very close and the 95% credible interval for the parameter r (logarithmic transformation model) contain only values larger than 1, the use of "model 2" could be a conservative better alternative for the data analysis of this data set (see also Figure 2).

	β		θ		r		
	Mean	SD.	Mean	SD.	Mean	SD.	Bayes Factor
Model 1	-1.810	1.258	0.296	0.300	—	—	$V_1 = \exp(-153.3)$
Model 2	-1.663	1.107	0.315	0.305	0.333	0.603	$V_2 = \exp(-156.3)$
Model 3	-1.568	1.144	0.338	0.320	—	—	$V_3 = \exp(-162.8)$

Table 2: Posterior Summaries Assuming the Transformation Models (Example 2)

8.3. Example 3

In this example, the data refer to a study described, carried at with 90 males' patients diagnosed in the period 1970-1978 with laryngeal cancer and who were followed up to 01/01/1983 [4]. For each patient, the age (in year) was recorded at diagnosis and the stage of the disease (I = primary tumour, II = involvement of nodules, III = metastasis, IV = combines of three previous stages) as well as their respective failure or censor times (in months). The stages are ordered by the degree of seriousness of the disease (less serious to more serious).

Assuming the proportional hazards model where $h(t \mid z)h_0(t) e^{\beta z}$ with covariate vector $\mathbf{z} = (z_1,z_2,z_3,z_4,z_5,z_6,z_7)$, z_1 denotes stage II, z_2 denotes stage III, z_3 denotes stage IV, z_4 denotes age, z_5 denotes the interaction between age and stage II, z_6 denotes the interaction between age and stage III, and z_7 denotes interaction between age and stage IV, we obtained the maximin likelihood estimator (MLE) for the vector of regression parameters β (use of the software R), given b $\widehat{\beta}_1 = -7.9461, \widehat{\beta}_2 = -0.1225, \widehat{\beta}_3 = 0.8470, \widehat{\beta}_4 = -0.0026, \widehat{\beta}_5 = 0.1203, \widehat{\beta}_6 = 0.0114$, and $\widehat{\beta}_7 = 0.0137$. [1].

For a Bayesian analysis of the data considering the semiparametric models introduced in section 2 (the proportional hazards model denoted as "model 1", the logarithmic transformation model denoted as "model 2" and the proportional odds ratio model ($r = 1$) denoted as "model 3" in the logarithmic transformation model) we also assume that the hazard function $h_0(t_i), i = 1,2,\cdots,n$ is an unknown latent factor, with a gamma distribution $G(0.1,0.1)$.

Assuming "model 1" (PH model), we assume gamma G(a,b) prior distributions for the parameters $\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6$, and θ_7, that is, $\theta_1 = \exp(\beta_1) \sim G(354,1000000); \theta_2 = \exp(\beta_2) \sim G(885,1000); \theta_3 = \exp(\beta_3) \sim G(2332,1000); \theta_4 = \exp(\beta_4) \sim G(997,1000); \theta_5 = \exp(\beta_5) \sim G(1128,1000); \theta_6 = \exp(\beta_6) \sim G(1011,1000)$ and $\theta_7 = \exp(\beta_7) \sim G(1014,1000)$.

In the elicitation of these prior distributions we used the prior information of the Cox PH model where β_1 was estimated by -7.9461, that is, $\exp(-7.9461) = 0.000354$; β_2 was estimated by -0.1225, that is, $\exp(-0.1225) = 0.884706$; β_3 was estimated by 0.8470, that is, $\exp(0.8470) = 2.33264$; β_4 was estimated by -0.0026, that is, $\exp(-0.0026) = 0{,}997403$; β_5 was estimated by 0.1203, that is, $\exp(0.1203) = 1.12784$; β_6 was estimated by 0.0114, that is, $\exp(0.0114) = 1.01147$ and β_7 was estimated by 0.0137, that is, $\exp(0.0137) = 1.01379$. We also assume the same gamma prior distributions for the parameters $\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,\theta_6$, and θ_7 considered for "model 1" (PH model) assuming "model 2" (logarithmic transformation model) and "model 3" (PO model) and a gamma $G(1,1)$ prior distribution for the parameter r in "model 2" (logarithmic transformation model).

For all cases, we used the OpenBUGS software considering a burn-in sample of 31,000 simulated samples discarded to eliminate the effects of the initial values in the iterative procedure and taking a final sample of size 1,000 (every 50*th* in 50,000 generated Gibbs samples) to get the Monte Carlo Carlo estimates for the parameters of interest [43]. The convergence of the Gibbs sampling algorithms was verified from trace plots of the simulated samples for each parameter. Table 3 shows the posterior summaries for each assumed model.

		Mean	SD.	95% credible interval	Bayes Factor
Model 1					
	β_1	-7.945	0.054	(-8.048 ; -7.844)	
	β_2	-0.122	0.033	(-0.189 ; -0.058)	
	β_3	0.846	0.020	(0.806 ; 0.886)	
	β_4	-0.020	0.007	(-0.033 ; -0.004)	$V_1 = -113.0$
	β_5	0.113	0.010	(0.094 ; 0.132)	
	β_6	0.016	0.013	(-0.008 ; 0.045)	
	β_7	0.016	0.015	(-0.012 ; 0.051)	
Model 2					
	β_1	-7.948	0.053	(-8.058 ; -7.850)	
	β_2	-0.103	0.033	(-0.161 ; -0.039)	
	β_3	0.850	0.012	(0.826 ; 0.871)	
	β_4	-0.011	0.010	(-0.029 ; 0.008)	$V_2 = -135.8$
	β_5	0.126	0.019	(0.094 ; 0.166)	
	β_6	0.017	0.015	(-0.007 ; 0.047)	
	β_7	0.023	0.013	(0.006 ; 0.049)	
	r	1.341	0.020	(1.300 ; 1.380)	
Model 3					
	β_1	-7.942	0.053	(-8.053 ; -7.839)	
	β_2	-0.124	0.033	(-0.191 ; -0.061)	
	β_3	0.847	0.020	(0.807 ; 0.888)	
	β_4	-0.018	0.011	(-0.041 ; 0.003)	$V_3 = -130.7$
	β_5	0.127	0.013	(0.096 ; 0.155)	
	β_6	0.021	0.015	(-0.007 ; 0.052)	
	β_7	0.023	0.018	(-0.012 ; 0.062)	

Table 3: Posterior Summaries Assuming the Transformation Models (Example 3)

From the results of Table 3, we observe that "model 1" (the PH model) is the best fitted model for the data using the posterior Bayes factor as a discrimination criterion, since the Bayesian Monte Carlo estimates for $V_k, k = 1,2,3$ in (13) are given respectively by, $V_1 = \exp(-113.0), V_2 = \exp(-135.8)$ and $V_3 = \exp(-130.7)$, with larger value for V_1, an indication that "model 1" (PH model) is the best model fitted by the data set. Assuming "model 1" (the PH model), we observe that the covariates age and interaction between age and stage III do not present significant effects on the response survival time since the 95% credible intervals for the regression parameters β_4 and β_6 contain the zero value. All the other covariates have significant effects on the response of interest.

Figure 3 shows the plots of the Kaplan-Meier nonparametric estimates for the survival functions considering each categorized covariate. From the plots of Figure 3 we observe that the needed assumption of PH Cox model (not crossing curves considering the categorized covariates) is observed considering each categorized covariate.

8.4. Example 4

In this example, let us consider the survival times (T) in days and a set of prognostic factors or covariates from 137 lung cancer patients, presented in appendix I of [7]. The covariates include the Karnofsky measure of the overall performance status (KPS) of the patients at into the trial, time in months from diagnosis to entry into the trial (DIAGTIME), age in year (AGE), prior therapy (INDPRI, yes or no), histological type of tumour, and type the therapy. There are four histological types of tumour: adeno, small, large, and squamous cell and two types of therapies: standard and experimental. The value of KPS have the following meanings: 10 – 30 completely hospitalized, 40 – 60 partial confinement, 70 – 90 able to care to self.

First, we define several index (or dummy) variables for the categorical variables and the censoring status. Let CENS = 0 when the survival time T is censored and 1 otherwise; IDADE = 1, INDSMA = 1 and INDSQU = 1 if the type of cancer cell is adeno, small, and squamous, respectively, and 0 otherwise; INDTHE = 1 if the standard therapy is received and 0 otherwise; and INDPRI = 1 if there is a prior therapy and 0 otherwise.

In this application, we assume only the covariate KPS and three cancer cell index variables: INDSQU, INDADE and INDSMA. First assuming the Cox PH model (1) we obtained from the partial likelihood using the R software the MLE estimates: -0.0229 (0.0044) for the regression parameter β_1 (KPS); -0.0814 (0.2794) for the regression parameter β_2 (INDSQU); 1.0610 (0.2950) for the regression parameter β_3 (INDADE) and 0.6940 (0.2532) for the regression parameter β_4 (INDSMA). Only the covariate INDSQU do not show significant effect on the survival times (p-value > 0.05).

For a Bayesian analysis of the data considering the semiparametric models introduced in section 2 (the proportional hazards model denoted as "model 1", the logarithmic transformation model denoted as "model 2" and the proportional odds ratio model (r = 1) denoted as "model 3" in the logarithmic transformation model) we also assume that the hazard function $h_0(t_i), i = 1,2,\ldots,n$ is an unknown latent factor, with a gamma distribution $G(0.1,0.1)$. We also assume gamma G(a,b) prior *distributions* for the parameters $\theta_1,\theta_2,\theta_3$, and θ_4, that is, $\theta_1 = \exp(\beta_1) \sim G(9.77,10);\theta_2 = \exp(\beta_2) \sim G(9.22,10);\theta_3 = \exp(\beta_3) \sim G(28.89,10)$ and $\theta_4 = \exp(\beta_4) \sim G(20.02,10)$.

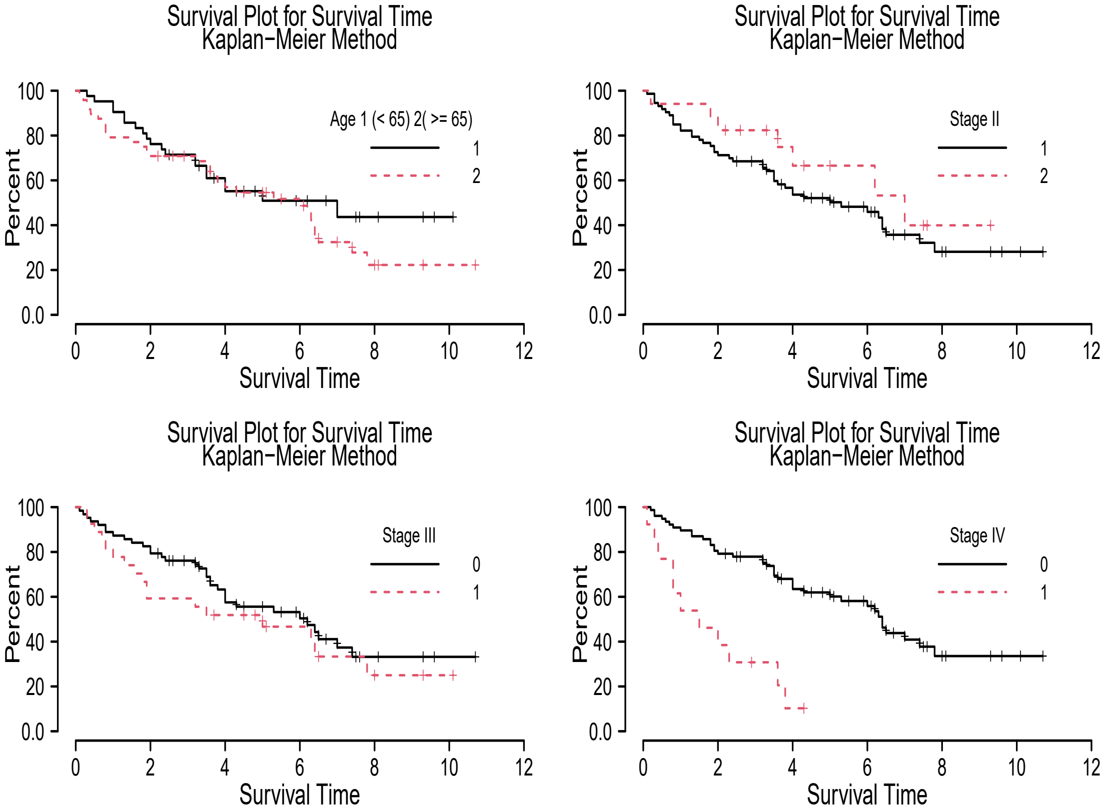

Figure 3: Kaplan-Meier Estimated Curves Considering Each Covariate (Example 3)

In the elicitation of the prior distributions we used the prior information of the Cox PH model (1) where β_1 was estimated by -0.022981, that is, $\exp(-0.022981) = 0.977281;\beta_2$ was estimated by -0.081376, that is, $\exp(-0.081376) = 0.921847;\beta_3$ was estimated by 1.061042, that is, $\exp(1.061042) = 2.889379$ and β_4 was estimated by 0.694003, that is, $\exp(0.694003) = 2.001713$. We also assume the same gamma prior distributions for the parameters $\theta_1,\theta_2,\theta_3$ and θ_4 considered for "model 1" (PH model) assuming "model 2" (logarithmic transformation model) and "model 3" (PO model) and a gamma G(1,1) prior distribution for the parameter r in "model 2" (logarithmic transformation model). For all cases, we have used the OpenBUGS software considering a burn-in sample of 11,000 simulated samples discarded to eliminate the effects of the initial values in the iterative procedure and taking a final sample of size 1,000 (every 100^{th} in 100,000 generated Gibbs samples) to get the Monte Carlo estimates for the parameters of interest [43]. The convergence of the Gibbs sampling algorithms was verified from trace plots of the simulated samples for each parameter. Table 4 shows the posterior summaries for each assumed model.

From the results of Table 4, we observe that "model 1" (the PH model) is the best fitted model for the data using the posterior Bayes factor as a discrimination criterion, since the Bayesian Monte Carlo estimates for $V_k, k = 1.2.3$ in (13) are given by $V_1 = \exp(-715.8), V2 = \exp(-725.4)$ and $V_3 = \exp(-762.1)$, with larger value for V_1, an indication that "model 1" (PH model) is the best model fitted by the data set. Assuming "model 1" (the

PH model), we observe that the covariates KPS, INDADE and INDSMA show significant effects on the survival times since the 95% credible intervals for the regression parameters β_1,β_3 and β_4 do not contain the zero value.

		Mean	SD.	95% Credible Interval	Bayes Factor
Model 1					
	β_1	-0.066	0.005	(-0.076 ; -0.056)	
	β_2	-0.0044	0.299	(-0.660 ; 0.493)	$V_1 = -715.8$
	β_3	0.996	0.188	(0.613 ; 1.371)	
	β_4	0.640	0.210	(0.226 ; 1.030)	
Model 2					
	β_1	-0.065	0.005	(-0.075 ; -0.053)	
	β_2	-0.050	0.308	(-0.689 ; 0.524)	
	β_3	1.010	0.185	(0.629 ; 1.353)	$V_2 = -725.4$
	β_4	0.661	0.210	(0.247 ; 1.077)	
	r	0.171	0.153	(0.006 ; 0.579)	
Model 3					
	β_1	-0.054	0.006	(-0.066 ; -0.041)	
	β_2	-0.046	0.288	(-0.631 ; 0.481)	$V_3 = -762.1$
	β_3	1.016	0.193	(0.611 ; 1.366)	
	β_4	0.654	0.218	(0.222 ; 1.072)	

Table 4: Posterior Summaries Assuming the Transformation Models (Example 4)

Figure 4 shows the plots of the Kaplan-Meier nonparametric estimates for the survival functions considering each categorized covariate. From the plots of Figure 4 we observe that the needed assumption of PH Cox model (not crossing curves) are observed in this example in agreement with the obtained results. Also observe from the results in Table 4, that the 95% credible interval for the parameter r (logarithmic transformation model or "model 2") is given by (0.006;0.579) an indication that the PH model is the best model (values close to zero).

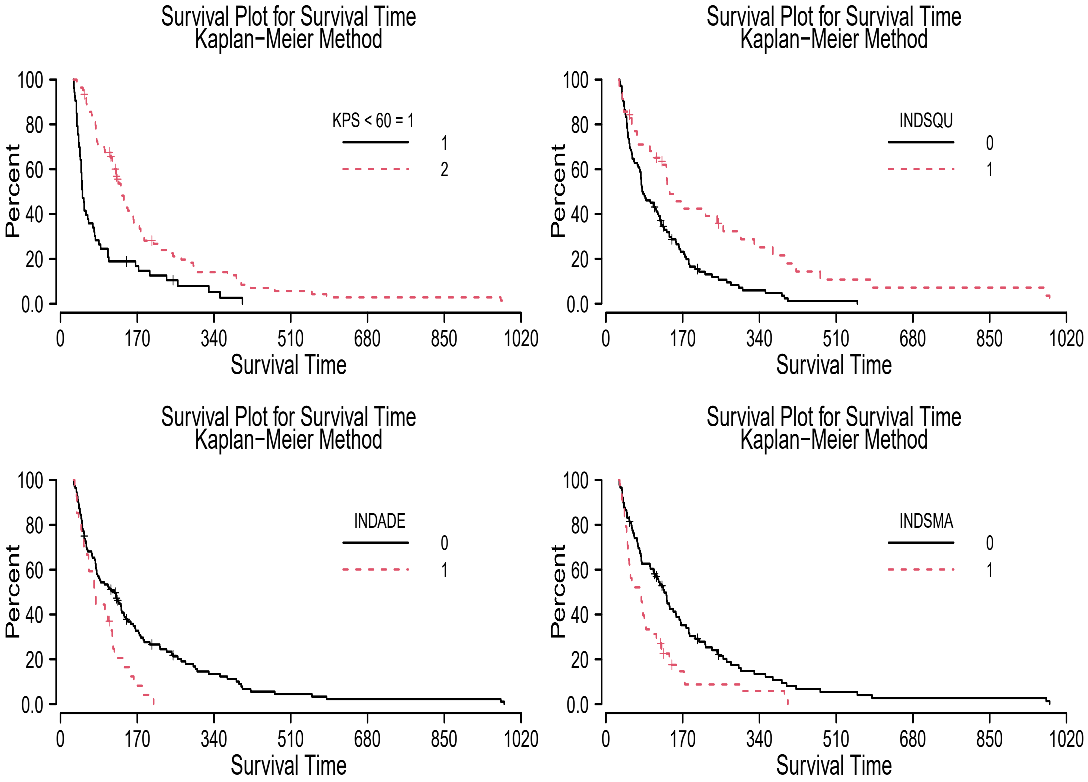

Figure 4: Kaplan-Meier Estimated Curves Considering Each Covariate (Example 4)

9. A Simulation Study

In this section, we assume simulated sample from the Weibull distribution with probability density function (pdf) $f(t) = at^{a-1}e^{-(t/b)a}/b^a, t > 0$ with survival function $S(t) = e^{-(t/b)a}$ and hazard function $h(t) = f(t)/S(t) = at^{a-1}/b^a$, where a is the shape parameter and b is the scale parameter. Let us assume only a covariate Z ($Z = 0$ for treatment 1 or control and $Z = 1$ for treatment 2) and the proportional hazard model $h(t) = h_0(t)e^{\beta z}$. For treatment 1 (control $Z = 0$) we have hazard function $h_1(t) = at^{a-1}/(b_1)^a$ and for treatment 2 ($Z = 1$), we have hazard function $h_2(t) = at^{a-1}/(b_2)a$, that is, we are assuming the same shape parameter a, but different scale parameters b_1 and b_2. That is, $h_2(t) = at^{a-1}/(b_2)^a = (b_1^a/b_2^a)at^{a-1}/(b_1)^a = e^{\beta}at^{a-1}/(b_1)^a = e^{\beta}h_1(t)$ (proportional hazard), where $e^{\beta} = (b_1/b_2)^a$, or, $\beta = \mathrm{a}\log(b_1/b_2)$. For simulation study, we assume $a = 1.5, b_1 = 30, b_2 = 50$, that is, $\beta = a\log(b_1/b_2) = 1.5\log(30/50) = -0.766238$. Also, let us consider $n = 100$. Table 5 shows the generated data (without censoring and with censoring). The five groups of simulated data are given by: group 1 with 100 observations without censoring, group 2 with 100 observations in presence of a small proportion of randomly censored (8 censored observations), group 3 with 100 observations in presence of a moderate proportion of randomly censored (27 censored observations), group 4 with 100 observations in presence of a large proportion of randomly censored (43 censored observations) and group 5 with 100 type I censoring – observations censored for > 50, that is, censored observations denoted by 50+ with 13 censored observations).

For a Bayesian analysis of the data considering the semiparametric models introduced in section 2 (the proportional hazards model denoted as "model 1", the logarithmic transformation model denoted as "model 2" and the proportional odds ratio model (r = 1) denoted as "model 3" in the logarithmic transformation model) we also assume that the hazard function $h_0(t_i), i = 1,2,\cdots,n$ is an unknown latent factor, with a gamma distribution $G(0.1,0.1)$. We assume a gamma $G(21.46,46)$ prior distribution for the parameter $\theta = \exp(\beta)$. Observe that $\theta = \exp(-0.766238) = 0.46652174 = 21.46/46$ (the mean of the gamma prior distribution for θ). Table 5 shows the posterior summaries of interest. Assuming the PH model and the partial likelihood function proposed the maximum likelihood estimate of the regression parameter (data without censored observations) is given by - 0.862(0.422) [1].

From the obtained Bayesian inference results using standard existing MCMC methods, we observe from Table 5 the robustness of the proposed methodology, considering different proportions of censored data. I all cases the posterior Bayesian factor criterion indicates the PH model, a special case of the semiparametric or transformation model, as the best model (larger value of V1) to be assumed in the data analysis (correct decision).

Group 1 (without censoring)									
2.9	52.7	21.6	18.8	20.2	65.3	43.1	44.5	20.4	48.2
	(l)	(s)	(l)	(m)		(m)	(m,l)	(l)	(s)
29.8	16.5	46.4	2.7	44.4	34.7	20.6	42.9	41.6	18.2
	(l)	(l)	(l)			(s,m,l)	(s)	(s,l)	
31.2	1.2	7.2	17.9	47.6	5.1	44.7	24.1	7.7	34.0
			(m)	(l)	(l)	(m,l)			
16.3	11.2	16.4	74.1	12.3	3.6	29.0	59.8	4.8	39.3
	(l)		(s)		(l)		(l)	(l)	(s)
10.2	1.2	28.7	9.2	56.2	16.5	16.4	37.4	26.1	2.9
(m,l)	(l)	(m)	(m)	(m,l)	(m,l)		(m)		(m)
Group 2 (without censoring)									
38.4	71.4	14.0	59.9	29.9	25.4	21.7	49.8	24.9	80.9
	(m)	(l)	(l)	(l)			(l)		
34.3	75.1	78.1	26.5	54.5	16.9	95.9	47.6	0.9	13.2
	(l)	(l)	(m)	(m)	(m)	(m)	(l)		(m)
56.1	76.2	17.2	23.7	67.9	41.9	103.7	15.6	31.5	14.6
			(l)			(l)	(l)		(l)
30.0	53.0	30.5	69.6	77.0	73.8	47.1	22.0	57.9	110.9
(l)		(m,l)	(m)	(l)		(l)	(l)		(m,l)

59.9	10.7	149.9	32.7	19.5	37.1	41.5	29.9	29.9	28.7
(m,l)	(s,m)			(l)	(m,l)	(l)		(m)	(m)
Type I censoring – observations $>$ 50 are censored in 50+									
All observations above (without censoring)but considering observations $>$ 50 as censored in 50.									

(s) - small proportion of 8 censored observations;

(m) - moderate proportion of 27 censored observations;

(l) - large proportion of 43 censored observations

Table 5: Simulated Data (n = 100; a = 1.5, b_1 = 30, b_2 = 50, β = −0.766238)

	β	θ	r	Bayes Factor
Data without censoring				
Model 1	-0.8204(0.2184)	0.4507(0.0970)	—	$V_1 = -482.3$
Model 2	-0.8082(0.2096)	0.4555(0.0959)	0.6137(0.4194)	$V_2 = -508.4$
Model 3	-0.8011(0.2191)	0.4596(0.0999)	—	$V_3 = -523.9$
Small Prop censoring				
Model 1	-0.7963(0.2235)	0.4622(0.1011)	—	$V_1 = -422.6$
Model 2	-0.8052(0.2281)	0.4586(0.1034)	0.5804(0.4057)	$V_2 = -465.6$
Model 3	-0.7893(0.2150)	0.4647(0.0997)	—	$V_3 = -481.1$
Moderate Prop censoring				
Model 1	-0.7990(0.2121)	0.4599(0.0973)	—	$V_1 = -349.6$
Model 2	-0.7874(0.2261)	0.4666(0.1049)	0.6025(0.4776)	$V_2 = -367.6$
Model 3	-0.7970(0.2227)	0.4618(0.1015)	—	$V_3 = -379.7$
Large Prop censoring				
Model 1	-0.8036(0.2157)	0.4582(0.0990)	—	$V_1 = -279.7$
Model 2	-0.7887(0.2243)	0.4666(0.1020)	0.6230(0.4693)	$V_2 = -294.3$
Model 3	-0.8055(0.2196)	0.4575(0.0982)	—	$V_3 = -303.5$
Censoring in 50+				
Model 1	-0.8039(0.2202)	0.4583(0.0983)	—	$V_1 = -342.2$
Model 2	-0.7978(0.2126)	0.4603(0.0945)	0.5520(0.4285)	$V_2 = -360.1$
Model 3	-0.8000(0.2099)	0.4592(0.0954)	—	$V_3 = -373.7$

Table 6: Posterior Summaries (Simulated Data)

10. Concluding Remarks

From the results obtained in this study, we observed that the use of a Bayesian approach for the semiparametric models in presence of covariates and censored data considering the complete likelihood function obtained from semiparametric or transformation models, where the unknown hazards are assumed as non-observed latent variables, could be a good alternative to get the inferences of interest in medical applications. In the special case of the usual PH (proportional hazards) model proposed, possibly the most used statistical methodology in lifetime data analysis in medical applications, the use of the proposed methodology can be a good alternative for the use of the standard partial likelihood function proposed by [1].

The proposed generalized semiparametric class of models includes the most common models as the proportional hazards model, the logarithmic transformation model and the proportional odds ratio model (r = 1) in the logarithmic transformation model.

The elicitation of the prior distributions for the regression parameters in this study was based on prior information assuming initially the PH model in the data analysis where the inference results were obtained from the maximum likelihood estimates obtained using the partial likelihood function (use of empirical Bayesian methods). Other informative prior distributions also could be used assuming prior information of medical experts.

It is important to point out that in many applications, researchers, especially in medical area, decide on a subjective way, by the proportional hazards model or the proportional odds model as special cases, usually based on [45]) nonparametric estimates for the survival functions or using residual methods as the residuals proposed and modified to check if the assumed PH model was appropriate [47, 48]. In many applications we could be not comfortable that the choose model is appropriate in the statistical analysis of the lifetime data usually in presence of covariates and censored data. In our approach we used a Bayesian discrimination criterion based on the posterior Bayes factor introduced that could be a very useful and simple tool to be used to verify if the assumed class of semiparametric model is appropriate in the lifetime data analysis [42].

It is important to point out that the classical inference results (MLE, confidence intervals, standard errors of the estimators) are obtained from the partial likelihood function assuming the PH model (1), using asymptotic normality results, which in general depend on large sample sizes to obtain good accuracy.

The obtained results of this study could be of interest to medical researchers since the family of proportional hazards models used extensively in medical studies, and generalizations given by the semiparametric family or transformation models, could have many advantages when compared to many existing parametric lifetime regression models usually considering generalizations of the Weibull distribution with three or more parameters [49].

References

1. Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological), 34*(2), 187-202.
2. Cox, D. R. (1975). Partial likelihood. *Biometrika, 62*(2), 269-276.
3. Lawless, J. F. (1982). Statistical Model and Method for Lifetime Data. New York: John Willey and Sons.
4. Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (Vol. 1230). New York: Springer.
5. Cox, D. R., & Oakes, D. (1984). *Analysis of survival data* (Vol. 21). CRC press.
6. Bennett, S. (1983). Analysis of survival data by the proportional odds model. *Statistics in medicine, 2*(2), 273-277.
7. Kalbfleisch, J. D., & Prentice, R. L. (2011). *The statistical analysis of failure time data.* John Wiley & Sons.
8. Yang, S., & Prentice, R. (2005). Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika, 92*(1), 1-17.
9. Demarqui, F. N., Mayrink, V. D., & Ghosh, S. K. (2019). An Unified Semiparametric Approach to Model Lifetime Data with Crossing Survival Curves. *arXiv preprint arXiv:1910.04475.*
10. He, W., & Yi, G. Y. (2020). Parametric and semiparametric estimation methods for survival data under a flexible class of models. *Lifetime Data Analysis, 26*, 369-388.
11. Li, J., Yu, T., Lv, J., & Lee, M. L. T. (2021). Semiparametric model averaging prediction for lifetime data via hazards regression. *Journal of the Royal Statistical Society Series C: Applied Statistics, 70*(5), 1187-1209.
12. Race, J. A., & Pennell, M. L. (2021). Semi-parametric survival analysis via Dirichlet process mixtures of the First Hitting Time model. *Lifetime Data Analysis, 27*, 177-194.
13. Yang, L., & Niu, X. F. (2021). Semi-parametric models for longitudinal data analysis. *J Financ Econ, 9*, 93-105.
14. Zhou, Q., Hu, T., & Sun, J. (2017). A sieve semiparametric maximum likelihood approach for regression analysis of bivariate interval-censored failure time data. *Journal of the American Statistical Association, 112*(518), 664-672.
15. Selingerova, I., Katina, S., & Horova, I. (2021). Comparison of parametric and semiparametric survival regression models with kernel estimation. *Journal of Statistical Computation and Simulation, 91*(13), 2717-2739.
16. Guo, S., & Zeng, D. (2014). An overview of semiparametric models in survival analysis. *Journal of Statistical Planning and Inference, 151*, 1-16.
17. Lin, D. Y., & Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika, 81*(1), 61-71.
18. Lin, D. Y., Wei, L. J., Yang, I., & Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 62*(4), 711-730.
19. Lin, D. Y., Wei, L. J., & Ying, Z. (2001). Semiparametric transformation models for point processes. *Journal of the American Statistical Association, 96*(454), 620-628.
20. Ma, S., & Kosorok, M. R. (2005). Penalized log-likelihood estimation for partly linear transformation models with current status data.
21. Rossini, A. J., & Tsiatis, A. A. (1996). A semiparametric proportional odds regression model for the analysis of current status data. *Journal of the American Statistical Association, 91*(434), 713-721.
22. Song, X., Davidian, M., & Tsiatis, A. A. (2002). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics, 58*(4), 742-753.
23. Song, X., & Wang, C. Y. (2008). Semiparametric approaches for joint modeling of longitudinal and survival data with time-varying coefficients. *Biometrics, 64*(2), 557-566.
24. Zeng, D., & Cai, J. (2010). A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika, 97*(3), 699-712.
25. Zeng, D., Yin, G., & Ibrahim, J. G. (2005). Inference for a class of transformed hazards models. *Journal of the American Statistical Association, 100*(471), 1000-1008.
26. Zeng, D., Lin, D. Y., & Lin, X. (2008). Semiparametric transformation models with random effects for clustered

failure time data. *Statistica Sinica, 18*(1), 355.

27. Zeng, D., & Lin, D. Y. (2009). Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics, 65*(3), 746-752.
28. Zeng, D., Mao, L., & Lin, D. Y. (2016). Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika, 103*(2), 253-271.
29. Chen, L., Lin, D. Y., & Zeng, D. (2012). Checking semiparametric transformation models with censored data. *Biostatistics, 13*(1), 18-31.
30. Chen, K., Jin, Z., & Ying, Z. (2002). Semiparametric analysis of transformation models with censored data. *Biometrika, 89*(3), 659-668.
31. Sun, J., & Sun, L. (2005). Semiparametric linear transformation models for current status data. *Canadian Journal of Statistics, 33*(1), 85-96.
32. Gao, F., Zeng, D., & Lin, D. Y. (2018). Semiparametric regression analysis of interval-censored data with informative dropout. *Biometrics, 74*(4), 1213-1222.
33. Achcar, J. A., Barili, E., & Martinez, E. Z. (2023). Semiparametric transformation model: A hierarchical Bayesian approach. *Model Assisted Statistics and Applications, 18*(3), 245-256.
34. Achcar, J. A., & Barili, E. (2023). Semiparametric transformation model in presence of cure fraction: a hierarchical Bayesian approach assuming the unknown hazards as latent factors. *Statistical Methods & Applications,* 1-24.
35. Abramowitz, M., Stegun, I. A., & Romer, R. H. (1988). Handbook of mathematical functions with formulas, graphs, and mathematical tables.
36. Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Texts in statistical science: Bayesian data analysis.
37. Gelfand, A. E., & Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association, 85*(410), 398-409.
38. Chib, S., & Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician, 49*(4), 327-335.
39. Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464.
40. Carlin, B. P., & Louis, T. A. (1997). Bayes and empirical Bayes methods for data analysis.
41. Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 64*(4), 583-639.
42. Aitkin, M. (1991). Posterior bayes factors. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 53*(1), 111-128.
43. Spiegelhalter, D. J., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS version 1.4 user manual. *MRC Biostatistics Unit, Cambridge. URL http://www.mrc-bsu.cam.ac. uk/bugs, 54.*
44. Acute Leukemia Group B, FREIREICH, E. J., GEHAN, E., FREI III, E. M. I. L., SCHROEDER, L. R., WOLMAN, I. J., ... & LEE, S. (1963). The effect of 6-mercaptopurine on the duration of steroid-induced remissions in acute leukemia: A model for evaluation of other potentially useful therapy. *Blood, 21*(6), 699-716.
45. Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association, 53*(282), 457-481.
46. Fleming, J. J., Parapia, L., Morgan, D. B., & Child, J. A. (1980). Increased Urinary B2-Microglobulin After Cancer Chemotherapy 1, 2. *Cancer Treatment Reports, 64*(4-5), 581-588.
47. Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika, 69*(1), 239-241.
48. Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*(3), 515-526.
49. Lai, C. D., & Lai, C. D. (2014). *Generalized weibull distributions* (pp. 23-75). Springer Berlin Heidelberg.