Brazilian Journal of Biometrics, **43**, 1–23 (2025)
doi:10.28951/bjb.v43i3.722

ARTICLE

Survival analysis of critically ill patients with cancer: use of semiparametric (transformation models) under a hierarchical Bayesian approach

Emerson Barili*[,1] and Jorge Alberto Achcar[2]

[1]Department of Statistics, State University of Maringá, Maringá-PR, Brazil
[2]Department of Social Medicine, Ribeirão Preto Medicine School, University of São Paulo, Ribeirão Preto-SP, Brazil
*Corresponding author. Email: ebarili2@gmail.com

(Received: February 7, 2024; Revised: March 20, 2025; Accepted: April 8, 2025; Published: July 28, 2025)

Abstract

Survival medical data in presence of covariates and censored data usually are analyzed assuming non-parametric or parametric regression modeling approaches as the popular proportional hazards models, the proportional odds models and the accelerated failure time models. In medical studies, it is usual the use of the popular proportional hazards models introduced by Cox, 1972 in the data analysis. Maximum likelihood estimation methods assuming the partial likelihood function introduced by Cox, 1975 are used to get the inferences of interest. In many applications, the assumption of proportional hazards could be non-verified which makes the use of the Cox model unfeasible. In this way, the use of semiparametric or transformation models recently introduced in the literature could be a good alternative in the analysis of lifetime data in presence of censoring and covariates. This class of models generalizes the popular class of proportional hazards models proposed by Cox, 1972 without the need to assume a parametric probability distribution for the survival times. In this study, we present a hierarchical Bayesian analysis considering semiparametric models to a data set consisting of the survival times of cancer patients admitted to the intensive treatment unit of the INCA health institute (Instituto Nacional de Câncer - INCA) in Rio de Janeiro, Brazil. The posterior summaries of interest are obtained using existing MCMC (Markov Chain Monte Carlo) simulation methods.

Keywords: semiparametric models; censored data; covariates; hierarchical Bayesian analysis; MCMC methods; cancer survival times.

1. Introduction

In the statistical analysis of medical survival data (time until the occurrence of a event of interest, for example, death) in presence of censored observations and covariates (Lawless, 1982; Klein &

Moeschberger, 1997; Cox & Oakes, 1984), usually medical researchers use standard non-parametric techniques, such as Kaplan & Meier, 1958 estimators for the survival function, log-rank or Wilcoxon non-parametric tests in the comparison of two or more treatments (Bradburn *et al.,* 2003) and the popular proportional hazards (PH) model (Cox, 1972), although this class of semiparametric models could be not suitable in many cases, with crossing survival curves. Other possibility is the use of proportional odds (PO) models (Bennett, 1983). As alternative to non-parametric methods, not much common in medical applications, we could use some standard parametric regression models (Kalbfleisch & Prentice, 2002) assuming some existing lifetime probability distributions as the exponential, Weibull, log-normal, gamma probability distributions or generalizations of these probability distributions (Bradburn *et al.,* 2003).

The literature introduced some generalized forms of the semipametric models including the PH and PO models to be used in lifetime data analysis as the semiparametric two sample strategy (Yang & Prentice, 2005) denoted as YP model and an unified approach introduced by Demarqui *et al.,* 2019 using Bernstein polynomials to model the baseline unknown hazard under both the frequentist and Bayesian frameworks. Li *et al.,* 2021 introduced a semiparametric model averaging prediction (SMAP) method which approximates the underlying unstructured nonparametric regression function by a weighted sum of low-dimensional nonparametric submodels. Race & Pennell, 2021 introduced a semi-parametric survival analysis via Dirichlet process mixtures of the First Hitting Time model considering several random effects specifications of the FHT model under a Bayesian approach. Yang & Niu, 2021 introduced semi-parametric models for longitudinal data analysis. Zhou *et al.,* 2017 considered semiparametric transformation models for interval-censored data. The use of semi-parametric models is becoming of great interest in different areas of applications.In this direction Ramos *et al.,* 2024 introduced a study related to the power-law distribution in pieces under a semi-parametric approach with change point detection.

In this study, we consider the statistical analysis of a data set assuming semiparametric models (data set in (Carvalho *et al.,* 2019)) and analyzed by Soares *et al.,* 2006, related to the survival times of cancer patients admitted to the intensive treatment unit of the INCA health institute (Instituto Nacional de Câncer - INCA) in Rio de Janeiro, Brazil.

The main goal of this study is to present a hierarchical Bayesian analysis considering the INCA cancer survival times data set assuming semiparametric models where the elicitation of prior distributions for the regression parameters of the model and for the unknown baseline hazard function is based on prior information obtained using standard non-parametric methods in a preliminary data analysis, that is, using empirical Bayesian methods (Carlin & Louis, 2000).The use of empirical Bayesian methods is becoming very popular in applications, although the use of the data twice leading to great debate within the statistical community due to potential biases introduced by reusing data. Some recent studies have focused in empirical Bayesian methods: Armstrong *et al.,* 2022 introduces the construction of robust empirical Bayesian confidence intervals in a normal means problem; Achcar *et al.,* 2017 introduced a empirical Bayesian approach in the elicitation of prior distributions for the parameters of the generalized gamma distribution; Efron, 2024 introduced the concepts and methods for the Empirical Bayes approach.

From now, the paper is prepared as follows: section 2 introduces the proportional (PH) hazards model; section 3 presents the semiparametric transformation models; section 4 introduces the likelihood function; section 5 introduces a hierarchical Bayesian approach for the class of semiparametric models; section 6 presents a statistical analysis of a INCA cancer survival time under the class of semiparametric models; finally section 7 introduces some concluding remarks.

2. The proportional (PH) hazards model

The proportional hazards (PH) model introduced by Cox, 1972 is a semi-parametric model where the hazard function is given by,

$$h(t;\mathbf{z}) = h_0(t)e^{\boldsymbol{\beta}'\mathbf{z}} \tag{1}$$

where $t > 0$ denotes the lifetime of a patient, $h_0(t)$ is the baseline hazard function defined as the limit of the probability of a individual to failure in the interval $[t, t+\Delta t]$ given that this individual is surviving until time t, that is,

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

$\boldsymbol{\beta}$ a vector of regression coefficients and Z a vector of covariates associated to each patient.

Under model (1), the hazard ratio for two different individuals with fixed covariate vectors is constant, that is, not dependent on time. It is observed two multiplicative components in model (1): one is non-parametric and the other is parametric, that is, we have a semi-parametric model. The covariates affect the hazard function in a multiplicative way according to the factor $\exp(\boldsymbol{\beta}'\mathbf{z})$. Cox, 1972 proposed a likelihood function that does not depend on the baseline hazard function $h_0(t)$, thus allowing inferences on $\boldsymbol{\beta}$ without the need to specify $h_0(t)$.

Considering n individuals under study and k distinct moments of observation of the event of interest (e.g., deaths), such that $t_1 < t_2 < \cdots < t_k$, where $k \leq n$, the partial (or marginal) likelihood function, proposed by Cox, 1972 is given by,

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{e^{\boldsymbol{\beta}' z_i}}{\sum_{l \in \mathbb{R}_i} e^{\boldsymbol{\beta}' z_l}} \tag{3}$$

where $\mathbb{R}_i$ is the set of individuals in risk at time t_i, and z_i is the vector of explanatory variables (covariates) associated with the individual having the event of interest at time t_i. Cox, 1975 showed that, although (3) is not a likelihood function in the usual sense, under very general conditions, it is verified the usual properties of the maximum likelihood estimators as the usual asymptotic properties of likelihood-based inference. The likelihood function (3) proposed by Cox, 1975 does not depend on the unknown hazard function $h_0(t)$.

3. Semiparametric transformation models

A generalization of the popular PH and PO models is given by the class of semiparametric transformation models (Zeng & Lin, 2009).

Under the semiparametric transformation model, the cumulative hazard function for the survival time T conditional on the covariate vector Z is given by,

$$\Lambda(t;\mathbf{z}) = G\left\{\int_0^t e^{\boldsymbol{\beta}\mathbf{z}} h(u)du\right\} \tag{4}$$

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

That is,

$$\Lambda(t;\mathbf{z}) = G\left\{e^{\boldsymbol{\beta}\mathbf{z}}\Lambda_0(t)\right\} \tag{5}$$

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

Further generalizations of the semiparametric transformation models were also introduced in the literature. Chen & Lu, 2012 introduced semiparametric transformation models in presence of a

cure fraction. Gao *et al.,* 2018 and Zeng *et al.,* 2016 introduced semiparametric regression analysis for interval-censored data. Other generalizations of the semiparametric model (or transformation models) could be seen in the literature (Chen *et al.,* 2002; Chen & Lu, 2012; Sun & Sun, 2005).

Achcar *et al.,* 2023 presented a hierarchical Bayesian approach for semiparametric models (or transformation models) assuming the unknown hazards as latent factors for semiparametric models; Achcar & Barili, 2023 introduced a hierarchical Bayesian approach for semiparametric models (or transformation models) in presence of cure fraction. Achcar & Barili, 2024. introduced a Bayesian approach for semiparametric models considering multivariate lifetime data in presence of censoring and covariates.

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

(a) 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 hazard ratio, $h(t;\mathbf{z}_i)/h(t;\mathbf{z}_j) = e^{\beta\mathbf{z}_i}h_0(t)/e^{\beta\mathbf{z}_j}h_0(t) = e^{\beta\mathbf{z}_i}/e^{\beta\mathbf{z}_j}$ (does not depend on t, that is, we have a proportional hazards model).

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

(c) If $G(x) = \log(1 + rx)/r(r \geq 0)$, the logarithmic transformation family, with $G(x) = x$ if $r = 0$ and $G(x) = \log(1 + x)$ if $r = 1$ (Zeng *et al.,* 2016). In this case we have $\Lambda(t;\mathbf{z}) = \log\{1 + re^{\beta\mathbf{z}}\Lambda_0(t)\}/r$ and $S(t;\mathbf{z}) = exp[-\Lambda(t;\mathbf{z})] = exp\{-\log[1 + re^{\beta\mathbf{z}}\Lambda_0(t)]/r\} = 1/[1 + e^{\beta\mathbf{z}}\Lambda_0(t)]$.

Remarks:

- The accumulated hazard function $\Lambda(t)$ can be given by $\Lambda(t) = \int_0^t h(u)du = -\log[S(t)]$, that is, $S(t) = exp[-\Lambda(t)]$.
- The hazard function $h(t)$ can be given by $h(t) = d\Lambda(t)/dt = -S'(t)/S(t)$.
- The density function $f(t)$ can be given by $f(t) = dF(t)/dt$ where $F(t) = 1 - S(t)$.
- (Abramowitz & Stegun, 1968): $\log(1+x) \approx x - x^2/2 + x^3/3 - \cdots$. ($|x| \leq 1$ 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).

4. The likelihood function

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

$$L(\cdot) = \prod_{i=1}^{n} f(\mathbf{t}_i;\mathbf{z}_i)^{\delta_i} S(\mathbf{t}_i;\mathbf{z}_i)^{1-\delta_i} \tag{6}$$

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 (6)) based in a sample of size n is given by,

$$L(\cdot) = \prod_{i=1}^{n} h(\mathbf{t}_i;\mathbf{z}_i)^{\delta_i} S(\mathbf{t}_i;\mathbf{z}_i) \tag{7}$$

4.1 The transformation $G(x) = x$ (a proportional hazards model)

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

$$\begin{aligned} L(\cdot) &= \prod_{i=1}^{n} h(t_i,\mathbf{z}_i)^{\delta_i} S(t_i;\mathbf{z}_i) \\ &= \prod_{i=1}^{n} \left[e^{\beta z_i} h_0(t_i)\right]^{\delta_i} e^{-e^{\beta z_i}\Lambda_0(t_i)} \\ &= e^{\sum_{i=1}^{n}\delta_i \beta \mathbf{z}_i} \left\{\prod_{i=1}^{n}\left[h_0(t_i)\right]^{\delta_i}\right\} e^{-\sum_{i=1}^{n} e^{\beta z_i}\Lambda_0(t_i)} \end{aligned} \tag{8}$$

4.2 The logarithmic transformation family $G(x) = \log(1 + rx)/r$

In this case we have,

$$\Lambda(t;\mathbf{z}_i) = \log\left\{1 + re^{\beta z}\Lambda_0(t)\right\}/r \quad \text{and}$$

$$S(t,\mathbf{z}) = \exp(-\Lambda(t,\mathbf{z})) = \exp\left(-\frac{\log(1 + re^{\beta z_i}\Lambda_0(t))}{r}\right)$$

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,

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

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

From (6) the likelihood function based on the ith observation is given by,

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

A special case of (11) is obtained when r =1 (proportional odds model).

5. A Bayesian approach

The use of Bayesian methods is becoming very popular in medical and epidemiology studies (Martinez & Achcar, 2014). A hierarchical Bayesian approach (Gelman *et al.,* 2004) is considered for the statistical analysis of semiparametric models assuming the unknown baseline hazard $h_0(t)$ as a latent factor (a non observed random variable). That is, we assume $d_i = h_0(t_i)$ as a random effect with a uniform probability distribution $U(a,b)$, with known hyperparameters a and b. The cumulative hazard function is given by, $\Lambda_0(t_i) = d_i t_i$, since $h_0(t_i)$ is the derivative of the cumulative hazard function $\Lambda_0(t_i)$, that is, $h_0(t_i) = d\Lambda_0(t_i)/dt$. In presence of only one covariate, we 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 and $G(c, d)$ denotes a gamma probability distribution with mean c/d and variance c/d^2. 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 $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3, \ldots, \theta_k)$, $\theta_j = exp(\beta_j), j = 1, 2, \ldots, k$, we assume independent gamma prior distributions $\theta_j \sim G(c_j, d_j)$.

We use the Bayes rule 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 $\boldsymbol{\theta}$ the vector of parameters of a model describing the behavior of the data D, if $P(\boldsymbol{\theta})$, $P(\boldsymbol{\theta}/D)$, and $L(D/\boldsymbol{\theta})$ indicate, respectively, the prior, the posterior distributions of $\boldsymbol{\theta}$, and the likelihood function of the model, then $P(\boldsymbol{\theta}/D) \propto L(D/\boldsymbol{\theta})P(\boldsymbol{\theta})$.

We use standard MCMC (Markov Chain Monte Carlo) methods as the Gibbs sampling algorithm and the Metropolis-Hastings algorithm (Gelfand & Smith, 1990; Chib & Greenberg, 1995) to get the posterior summaries of interest.

In the discrimination of the different models to be used in each application, we use the DIC (Deviance Information Criterion) proposed by Spiegelhalter *et al.,* 2002, which is appropriate when we use MCMC methods to get the posterior summaries of interest.

The DIC criterion is widely used in Bayesian inference applied especially using Markov Chain Monte Carlo methods. Set the deviation (deviance) by:

$$D(\boldsymbol{\theta}) = -2 \ln L(\boldsymbol{\theta}) + C \tag{12}$$

where $\boldsymbol{\theta}$ is a vector of unknown model parameters; $L(\boldsymbol{\theta})$ is the likelihood and C is a constant not necessarily known in the comparison of two models. The DIC criterion defined by Spiegelhalter *et al.,* 2002 is given by,

$$DIC = D(\widehat{\boldsymbol{\theta}}) + 2_{P_D} \tag{13}$$

in which $D(\widehat{\boldsymbol{\theta}})$ is the deviation calculated from the posterior mean $\widehat{\boldsymbol{\theta}} = E(\boldsymbol{\theta} \mid y)$ and p_D is the effective number of parameters in the model, given by,

$$p_D = \overline{D} - D(\widehat{\boldsymbol{\theta}})$$

where

$$\overline{D} = E(D(\boldsymbol{\theta}) \mid y)$$

is the posterior average of the deviation that measures the goodness of fit of the data for each model. Smaller DIC values indicate better models and these values can be negative.

6. Statistical analysis of the INCA cancer data

We consider a data set related to the survival times of cancer patients admitted to the intensive treatment unit of the INCA health institute (Instituto Nacional de Câncer - INCA) in Rio de Janeiro, Brazil. The data were obtained from a cohort of 862 cancer patients (data set in (Carvalho *et al.,* 2019)) and analyzed by Soares *et al.,* 2006, whose main objective was to evaluate if some factors (covariates) have statistical significant effects on the survival times of the patients. The variables in the original data set are given by,

- T: Time from the first case monitored in days.
- Death status: death = 1, censored observation = 0.
- Gender: male = 1, female = 0.
- Age: in years.

- Tumor type: (1) Locus = localized solid; (2) Mtx = metastatic ; (3) Hemato = hematological.
- Desnutrition: 1 for recent weight loss > 10% or BMI < 18; 0 = no.
- Comorbidity: 1 for severe comorbidities present; 0 = absent.
- Leucopenia: 1 = for leukopenia present; 0 = absent.

Remarks: Comorbidity in the data set describes the simultaneous occurrence of two or more health problems in the same individual. Leukopenia is when the number of leukocytes, which are the blood's defense cells, is low.

The data set presents 499 uncensored observations and 363 right censored observations. The survival times are denoted by t (survival in days) and the variable δ (death status) denotes de censoring indicator ($\delta = 1$ if not censored, $\delta = 0$ if censored).

For the categorized covariates, Figure 1 shows the plots of the non-parametric Kaplan & Meier, 1958 estimates for the survival functions in each level of the covariates. Using standard log-rank or Wicoxon non-parametric tests it is observed that the covariates tumor type, desnutrition, comorbidity and leukopenia shown significant differences for the survival times (p-value < 0.05).

A first analysis is considered assuming the standard proportional (PH) hazards model of Cox (1) with $exp(\mathbf{\beta z}) = exp(\beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_7 z_7)$, where Z_1 denotes the covariate gender (1=male; 0=female); Z_2 denotes the covariate age (years); Z_3 denotes a dummy covariate for tumor locus (locus=1; 0 other); Z_4 denotes a dummy covariate for tumor hemato (hemato=1; 0 other); Z_5 denotes the covariate desnutrition (1 for recent weight loss > 10% or BMI < 18; 0 = no); Z_6 denotes the covariate comorbidity (1 for severe comorbidities present; 0 = absent) and Z_7 denotes the covariate leucopenia (1 = for leukopenia present; 0 = absent). The covariate Z_8 (Mtx or metastatic tumor) is considered as reference.

The MLE estimates of the regression parameters of model (1) obtained from the partial likelihood function (use of an iterative numerical procedure) and the R software, are presented in Table 1 from where it is also observed that the covariates Z_2 (age), Z_3 (tumor locus), Z_5 (desnutrition), Z_6 (comorbidity) and Z_7 (leucopenia) shown significant effects on the lifetimes of the patients (p-value < 0.05).

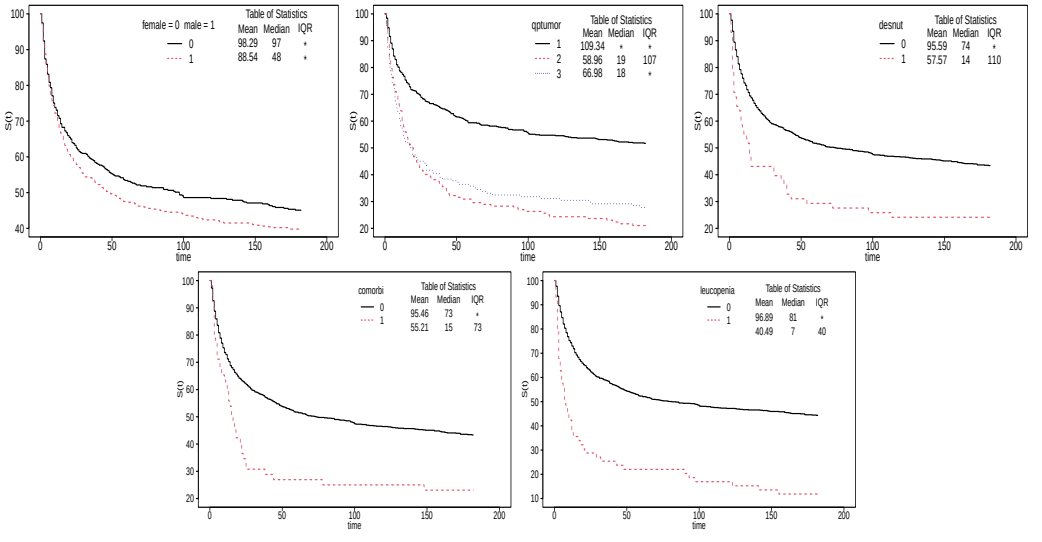


Figure 1. Kaplan-Meier (1958) estimates for the survival functions.

Table 1. Maximum likelihood estimates (MLE) of the PH Cox regression parameters

	coef	exp(coef)	se(coef)	z	p
sex	0.09916	1.10424	0.09125	1.09	0.27722
age	0.01306	1.01315	0.00298	4.38	< 0.001
tumorloco	-0.73584	0.47910	0.11456	-6.42	< 0.001
tumorhemato	-0.00729	0.99274	0.13682	-0.05	0.95751
desnut	0.55041	1.73397	0.16126	3.41	0.00064
comorbi	0.52255	1.68632	0.16769	3.12	0.00183
leucopenia	0.87309	2.39430	0.15838	5.51	< 0.001

We interpret the obtained results as follows: since the MLE of the regression parameter associated to age (Z_2) is positive, the hazard rate of deaths increases with higher ages; since the MLE of the regression parameter associated to tumor type locus (Z_3) is negative, the hazard rate of deaths decreases with tumor type locus in comparison to the other two types of tumor (result in agreement to the Kaplan-Meier plots in Figure 1); since the MLE of the regression parameter associated to desnutrition (Z_5) is positive, the hazard rate of deaths increases with the presence of desnutrition (result in agreement to the Kaplan-Meier plots in Figure 1); since the MLE of the regression parameter associated to comorbidity (Z_6) is positive, the hazard rate of deaths increases with the presence of comorbidity (result in agreement to the Kaplan-Meier plots in Figure 1); since the MLE of the regression parameter associated to leucopenia (Z_7) is positive, the hazard rate of deaths increases with the presence of leucopenia (result in agreement to the Kaplan-Meier plots in Figure 1).

To check if the proportional hazard assumption is verified, we could use plots of the weighted Schoenfeld residuals (Schoenfeld, 1982; Grambsch & Therneau, 1994) against survival times where the presence of some patterns in these graphs may indicate departures from the proportional hazards assumption. Figure 2 shows the plots of the Weighted Schoenfeld residuals (PH Cox regression model), from where it is difficult to say that the PH model is well fitted by the data set (a subjective decision). From the asymptotical chi-square test with one degree of freedom to test for the proportional hazards assumption with respect to each covariate (Grambsch & Therneau, 1994),we observed that for the covariate tumor hemato (p-value < 0.10) there is indication that the proportional hazards model is not apropriated although for all other covariates we obtained p-values > 0.10 (use of a 10% significance level).

As a second statistical analysis, we assume the semiparametric models introduced in section 3 with $exp(\boldsymbol{\beta z}) = exp(\beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_7 z_7)$, where Z_1 is gender (1=male; 0=female); Z_2 is age (years); Z_3 is tumor locus (locus=1; 0 other); Z_4 is tumor hemato (hemato=1; 0 other); Z_5 is desnutrition (1 for recent weight loss > 10% or BMI < 18; 0 = no); Z_6 is comorbidity (1 for severe comorbidities present; 0 = absent) and Z_7 is leucopenia (1 = for leukopenia present; 0 = absent), assuming a hierarchical Bayesian approach. The covariate Z_8 (Mtx or metastatic tumor) is considered as reference. From prior information obtained in Table 2 (use of empirical Bayesian methods, see Carlin & Louis, 1997) assuming the PH model using partial likelihood function, we assume the following prior gamma probability distributions for the reparameterized forms of the regression parameters $\theta_j = exp(\beta_j), j = 1, 2, \ldots, 7$ assuming the proportional (PH) hazards model denoted as "model 1" : $\theta_1 \sim G(110.424, 100); \theta_2 \sim G(101.315, 100); \theta_3 \sim G(47.910, 100); \theta_4 \sim G(99.274, 100); \theta_5 \sim G(173.397, 100); \theta_6 \sim G(168.632, 100)$ and $\theta_7 \sim G(239, 430, 100)$.

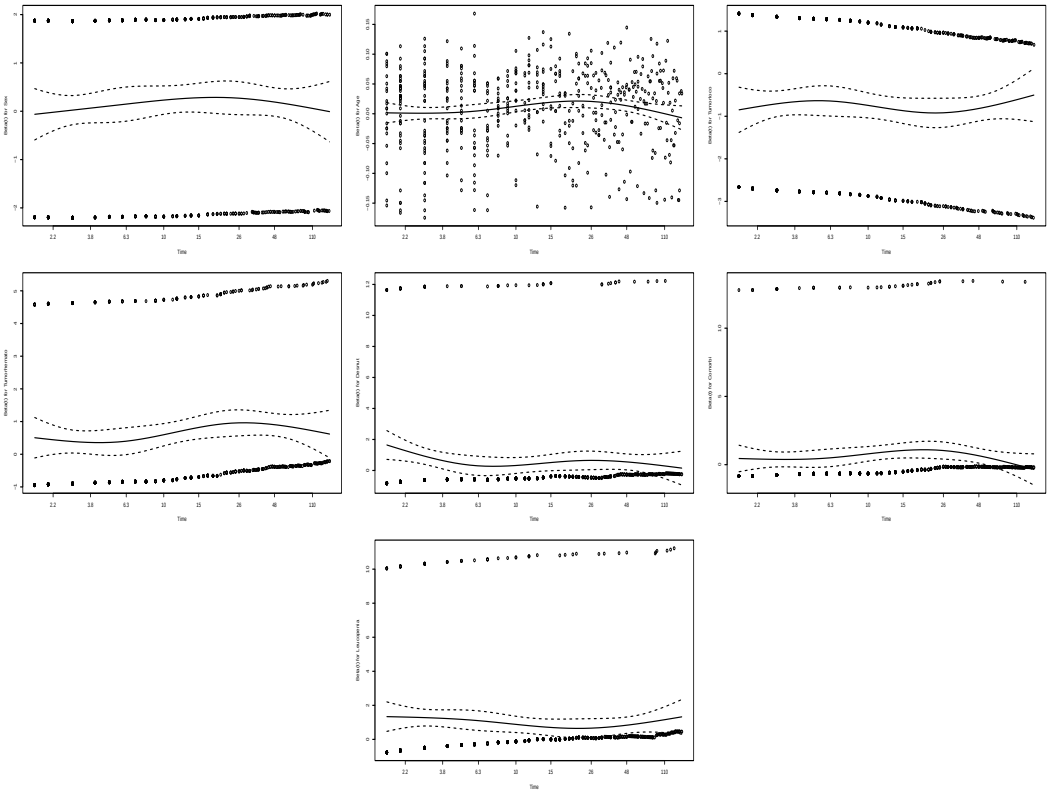


Figure 2. Weighted Schoenfeld residuals (PH Cox regression model).

We also assume the same gamma prior distributions for the parameters $\theta_j = exp(\beta_j), j = 1, 2, ..., 7$ assuming the logarithmic transformation family $G(x) = \log(1 + rx)/r$ or proportional odds model denoted as "model 2" and the special case of the logarithmic transformation family obtained when $r = 1$ denoted as "model 3" and a gamma $G(1, 1)$ prior distribution for the parameter r in "model 2".

Since the baseline hazard function is unknown in the semiparametric models introduced in section 2, we assume the hazard function $h_{0_i} = a_i, i = 1, 2, \ldots, 862$ as a random effect (a non-observed latent variable) with a uniform $U(0, 0.1)$ probability distribution. Observe that the accumulated hazard function is given by $\Lambda_0(t_i) = a_i t_i$, for $i = 1, 2, \ldots, 862$ where the hazard functions are given by $h_0(t) = d\Lambda_0(t)/dt$.

We used the OpenBUGS software (Spiegelhalter *et al.,* 2003), 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 Carlo estimates for the parameters of interest assuming the three models. The convergence of the Gibbs sampling algorithms was verified from trace plots of the simulated samples for each parameter and the Gelman & Rubin, 1992 index assuming three parallel chains (See Appendix 2 at the end of the manuscript).

Using DIC as a discrimination criterion, the Monte Carlo estimates of DIC (2002) are given, respectively by, DIC = 5281.0 ("model 1"), DIC = 5265.0 ("model 2") and DIC = 5215.0 ("model 3"), an indication that "model 3" based on the complete likelihood function with latent variables representing the unknown hazard functions is the best fitted model (better fitted model has smaller DIC).

Assuming "model 3" (logarithmic transformation family $G(x) = \log(1 + rx)/r$ or proportional

odds model with $r = 1$), Table 2 shows that the covariates Z_3 (tumor locus), Z_5 (desnutrition), Z_6 (comorbidity) and Z_7 (leucopenia) shown significant effects on the survival times of the patients since the 95% credible intervals for the corresponding regression parameters do not include the zero value.

It is interesting to observe that using the PH model under a classical approach (partial likelihood function proposed by Cox, 1972, the covariate Z_2 (age) also show significant effect (p-value < 0.05) but looking at the scatter plot of Figure 3 (times versus age considering only the uncensored observations) it is hard to say that there is some dependence between the survival times and the ages of the patients. That is, "model 3", a proportional odds model, under a Bayesian approach, is more realistic for the data set.

7. Concluding remarks

The class of semi-parametric models (transformation methods) recently introduced in the literature could be a good alternative in the analysis of survival data in presence of censored data and covariates, especially in medical applications. In medical studies, as observed in the medical literature, non-parametric techniques are commonly used in the lifetime data analysis, as the Kaplan-Meier non-parametric estimates of the survival function, log-rank and Wicoxon non-parametric tests to compare the survival curves and the proportional hazards model of Cox, 1972.

Table 2. Posterior summaries (mean, standard-deviation and 95% credible intervals) assuming "model 3"

	mean	sd	Lower 95% c.i	Upper 95% c.i
β_1	0.05943	0.07947	-0.1057	0.2098
β_2	-0.0019	0.0018	-0.0057	0.0015
β_3	-1.265	0.1189	-1.500	-1.043
β_4	-0.0469	0.0888	-0.232	0.1196
β_5	0.5944	0.0692	0.4604	0.7226
β_6	0.5577	0.0716	0.4137	0.6946
β_7	0.8893	0.0609	0.7680	1.007
θ_1	1.065	0.0845	0.8997	1.233
θ_2	0.9981	0.001837	0.9943	1.002
θ_3	0.2843	0.03366	0.2232	0.3524
θ_4	0.9579	0.08442	0.7929	1.127
θ_5	1.816	0.1258	1.585	2.06
θ_6	1.751	0.1253	1.512	2.003
θ_7	2.438	0.1483	2.155	2.736

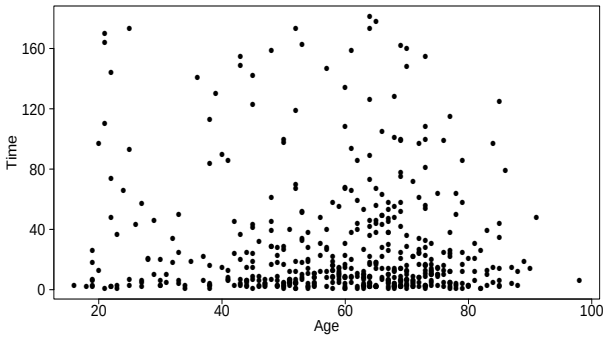


Figure 3. Scatter plots the survival functions (not censored) versus Z_2 (age).

In practice, the proportional hazards (PH) model (Cox, 1972) is the most used technique in the lifetime data analysis, although the assumption of proportional hazards could not be verified in many applications. Other common problem using the PH model under the partial likelihood function proposed by (Cox, 1972; Cox, 1975) is that the obtained inferences are based on asymptotical results usually requiring large sample sizes to get accurate results.

The class of semiparametric models (transformation models) generalizes the usual models assumed in medical applications such as the proportional hazards model or the odds ratio model, although under this class of models the baseline hazard function is unknown requiring to be approximated in some numerical way. In this study, we introduced a simple method to obtain inferences of interest assuming semiparametric models under a hierarchical Bayesian methodology where the unknown hazard functions are assumed as latent variables. The posterior summaries of interest are obtained from usual MCMC (Markov Chain Monte Carlo) methods assuming the elicitation of the prior distributions for the parameters of the model using some prior information obtained using the usual proportional hazards (PH) model of Cox, 1972.

Using the proposed methodology, we observed in an application to a data set consisting of the survival times of cancer patients admitted to the intensive treatment unit of the INCA health institute (Instituto Nacional de Câncer - INCA) in Rio de Janeiro, Brazil, more sensitive inferential results compared to the inferential results obtained using the traditional Cox proportional hazards model based on asymptotical normality of the maximum likelihood estimators of the regression parameters under the partial likelihood approach.

It is also important to point out that with the proposed methodology, using standard existing Bayesian discrimination methods as the DIC criterion, it was possible to choose the best class of semiparametric models, in our application, the odds ratio class, usually a not easy task in applications.

An alternative for the use of non-parametric or semi-parametric methods for the survival data analysis in presence of covariates and censored data could be the use of existing parametric regression models. Medical researchers argue that parametric regression models based on different probability distributions could present some difficulties in the choice of each parametric model for the survival data and difficulties in the medical interpretations for each parameter of the proposed parametric model assumed in each application. Considering standard Weibull and log-normal parametric regression models available in most statistical software in the analysis of the INCA cancer data set, the Cox & Snell, 1968 residual plots did not indicate good fit for the Weibull and log-normal regression models to the INCA cancer data set.

As a final remark, we can conclude that the hierarchical Bayesian approach for the generalized model (transformation model) introduced in this study could be of great interest when the assumption of proportional hazards model is not verified for the data. The proposed methodology is a useful complementation and generalization for the popular and powerful Cox, 1972 model, perhaps, the

most used statistical model in medical applications considering lifetime data in terms of efficiency, applicability and interpretability.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: ACHCAR, J.A.; BARILI, E. **Data curation**: ACHCAR, J.A.; BARILI, E. **Formal analysis**: ACHCAR, J.A.; BARILI, E. **Funding acquisition**: ACHCAR, J.A.; BARILI, E. **Investigation**: ACHCAR, J.A.; BARILI, E. **Methodology**: ACHCAR, J.A.; BARILI, E. **Project administration**: ACHCAR, J.A.; BARILI, E. **Software**: ACHCAR, J.A.; BARILI, E. **Resources**: ACHCAR, J.A.; BARILI, E. **Supervision**: ACHCAR, J.A.; BARILI, E. **Validation**: ACHCAR, J.A.; BARILI, E. **Visualization**: ACHCAR, J.A.; BARILI, E. **Writing - original draft**: ACHCAR, J.A.; BARILI, E. **Writing - review and editing.** ACHCAR, J.A.; BARILI, E.

References

[1] Abramowitz, M. & Stegun, I.A. Handbook of mathematical functions with formulas, graphs, and mathematical tables. *US Government printing office* **39**, (1968).

[2] Achcar, J. A. & Barili, E. Semiparametric transformation model in presence of cure fraction: a hierarchical Bayesian approach assuming the unknown hazards as latent factors. *Statistical Methods & Applications* **33**, 1–24 (2023). https://doi.org/10.1007/s10260-023-00734-w

[3] Achcar, J. A., Barili, E. & Martinez, E. Z. Semiparametric transformation model: A hierarchical Bayesian approach. *Model Assisted Statistics and Applications* **18**, 3, 245–256 (2023). https://doi.org/10.3233/MAS-221408

[4] Achcar, J.A. & Barili, E. Multivariate lifetime data in presence of censoring and covariates: Use of semiparametric models under a Bayesian approach. *Communications in Statistics-Theory and Methods* **33**, 1–30 (2024). https://doi.org/10.1080/03610926.2024.2400163

[5] Achcar, J. A., Ramos, P.L. & Martinez, E. Z. Some computational aspects to find accurate estimates for the parameters of the generalized gamma distribution. *Pesquisa Operacional* **37**, 2, 365-385 (2017). https://doi.org/10.1590/0101-7438.2017.037.02.0365

[6] Armstrong, T. B.; Kolesar, M. & Plagborg-Moller, M. Robust empirical Bayes confidence intervals. *Econometrica* **90**, 6, 2567-2602 (2022). https://doi.org/10.3982/ECTA18597

[7] Bennett, S. Analysis of survival data by the proportional odds model. *Statistics in Medicine* **2**, 2, 273–277 (1983). https://doi.org/10.1002/sim.4780020223

[8] Bradburn, M. J., Clark, T. G., Love, S. B. & Altman, D. G. Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. *British Journal of Cancer* **89**, 3, 431–436 (2003). https://doi.org/10.1038/sj.bjc.6601119

[9] Carlin, B.P. & Louis, T.A. Bayes and Empirical Bayes Methods for Data Analysis. *2nd Edition Chapman & Hall* (2000).

[10] Carvalho, M. S., Andreozzi, V. L., Codeço, C. T., Campos, D. P., Barbosa, M. T. S. & Shimakura, S. E. Análise de Sobrevivência: teoria e aplicações em saúde. *SciELO-Editora FIOCRUZ* (2019).

[11] Chen, C.-M. & Lu, T.-F. C. Marginal analysis of multivariate failure time data with a surviving fraction based on semiparametric transformation cure models. *Computational Statistics & Data Analysis* **56**, 1, 645-655 (2012). https://doi.org/10.1016/j.csda.2011.09.013

[12] Chen, K., Jin, Z. & Ying, Z. Semiparametric analysis of transformation models with censored data. *Biometrika* **89**, 3, 659–668 (2002). https://doi.org/10.1093/biomet/89.3.659

[13] Chib, S. & Greenberg, E. Understanding the metropolis-hastings algorithm. *The American Statistician* **49**, 4, 327–335 (1995). https://doi.org/10.2307/2684568

[14] Cox, D. R. Partial likelihood. *Biometrika*, **62**, 2, 269–276 (1975). https://doi.org/10.2307/2335362

[15] Cox, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* **34**, 2, 187–202 (1972). https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

[16] Cox, D. R. & Snell, E. J. A general definition of residuals. *Journal of the Royal Statistical Society: Series B (Methodological)*, **30**, 2, 248–265 (1968). https://doi.org/10.1111/j.2517-6161.1968.tb00724.x

[17] Cox, D. R. & Oakes, D. Analysis of survival data. *Chapman and Hall, New York* (1984). https://doi.org/10.1201/9781315137438

[18] Demarqui, F. N., Mayrink, V. D. & Ghosh, S. K. An Unified Semiparametric Approach to Model Lifetime Data with Crossing Survival Curves. *arXiv preprint arXiv:1910.04475* (2019). https://doi.org/10.48550/arXiv.1910.04475

[19] Efron, B. Empirical Bayes: Concepts and methods. In: Handbook of Bayesian, Fiducial, and Frequentist Inference. *Chapman and Hall/CRC* 8-34 (2024). https://doi.org/10.1201/9780429341731

[20] Gao, F., Zeng, D. & Lin, D.-Y. Semiparametric regression analysis of interval-censored data with informative dropout. *Biometrics* **74**, 4, 1213–1222 (2018). https://doi.org/10.1111/biom.12911

[21] Gelfand, A. E. & Smith, A. F. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 410, 398–409 (1990). https://doi.org/10.2307/2289776

[22] Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. Bayesian data analysis *Chapman and Hall/CRC* (2004).

[23] Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 4, 457–472 (1992). DOI: 10.1214/ss/1177011136

[24] Grambsch, P. M. & Therneau, T. M. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 3, 515–526 (1994). https://doi.org/10.2307/2337123

[25] Kalbfleisch, J. D. & Prentice, R. L. The statistical analysis of failure time data *John Wiley & Sons* (2002).

[26] Kaplan, E. L. & Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* **53**, 282, 457–481 (1958). https://doi.org/10.2307/2281868

[27] Klein, J. P. & Moeschberger, M. L. Survival analysis: Techniques for censored and truncated data. *Springer-Verlag, New York* (1997).

[28] Lawless, J. F. Statistical models and methods for lifetime data. *New York: Wiley* (1982).

[29] Li, J., Yu, T., Lv, J. & Lee, M.-L. T. Semiparametric model averaging prediction for lifetime data via hazards regression. *Journal of the Royal Statistical Society Series C: Applied Statistics* **70**, 5, 1187–1209 (2021). https://doi.org/10.1111/rssc.12502

[30] Martinez, E. Z. & Achcar, J. A. Trends in epidemiology in the 21st century: time to adoptvBayesian methods. *Cadernos de Saúde Pública* **30**, 4, 703–714 (2014). https://doi.org/10.1590/0102-311X00144013

[31] Race, J. A. & Pennell, M. L. Semi-parametric survival analysis via Dirichlet process mixtures of the First Hitting Time model. *Lifetime Data Analysis* **27**, 177–194 (2021). https://doi.org/10.1007/s10985-020-09514-0

[32] Ramos, P. L., Jerez-Lillo, N., Segovia, F. A., Egbon, O. A. & Louzada, F. Power-law distribution in pieces: a semi-parametric approach with change point detection. *Statistics and Computing***34**, 1, 16 (2024). https://doi.org/10.1007/s11222-023-10336-x

[33] Schoenfeld, D. Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 1, 239–241 (1982). https://doi.org/10.2307/2335876

[34] Soares, M., Carvalho, M. S., Salluh, J. I., Ferreira, C. G., Luiz, R. R., Rocco, J. R. & Spector, N. Effect of age on survival of critically ill patients with cancer. *Critical Care Medicine* **24**, 24 715–721 (2006). https://doi.org/10.1097/01.ccm.0000201883.05900.3f

[35] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **64**, 4, 583–639 (2002). https://doi.org/10.1111/1467-9868.00353

[36] Spiegelhalter, D. J., Thomas, A., Best, N. & Lunn, D. WinBUGS version 1.4 user manual. *MRC Biostatistics Unit, Cambridge.* **54** (2003). URL http://www. mrc-bsu. cam. ac. uk/bugs

[37] Sun, J. & Sun, L. Semiparametric linear transformation models for current status data. *Canadian Journal of Statistics* **33**, 1, 85–96 (2005). https://doi.org/10.1002/cjs.5540330107

[38] Yang, L. & Niu, X.-F. Semi-Parametric Models for Longitudinal Data Analysis. *Journal of Finances Economics* **9**, 93–105 (2021). DOI: 10.12691/jfe-9-3-1

[39] Yang, S. & Prentice, R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika* **92**, 1, 1–17 (2005). https://doi.org/10.1093/biomet/92.1.1

[40] Zeng, D. & Lin, D. Semiparametric transformation models with random effects for joint analysis of recurrent and terminal events. *Biometrics* **65**, 3, 746–752 (2009). https://doi.org/10.1111/j.1541-0420.2008.01126.x

[41] Zeng, D., Mao, L. & Lin, D. Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika* **103**, 2, 253–271 (2016). https://doi.org/10.1093/biomet/asw013

[42] Zhou, Q., Hu, T. & Sun, J. 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 (2017). https://doi.org/10.1080/01621459.2016.1158113

Appendix 1 (OpenBUGS codes)

Proportional hazards model

```
model {
for (i in 1:N) {
zeros[i] <- 0
phi[i] <- - loglike[i]
zeros[i] ~ dpois(phi[i])
loglike[i]<- delta[i]*(beta1*sex[i]+ beta2*age[i] + beta3*tumorloco[i] +
beta4*tumorhemato [i] +  beta5*desnut[i] + beta6*comorbi[i] +
beta7*leucopenia[i])+ delta[i]*log(lambda[i]) - exp(beta1*sex[i]+
beta2*age[i] + beta3*tumorloco[i] + beta4*tumorhemato [i] +
beta5*desnut[i] + beta6*comorbi[i] + beta7*leucopenia[i])*Lambda[i]
lambda[i] <-a[i]
Lambda[i] <- a[i]*t[i]
a[i]~ dunif(0,0.1)
}
beta1<- log(theta1)
theta1~ dgamma(110.424,100)
beta2<- log(theta2)
theta2~ dgamma(101.315,100)
beta3<- log(theta3)
theta3~ dgamma(47.910,100)
beta4<- log(theta4)
theta4~ dgamma(99.274,100)
beta5<- log(theta5)
theta5~ dgamma(173.397,100)
beta6<- log(theta6)
theta6~ dgamma(168.632,100)
beta7<- log(theta7)
theta7~ dgamma(239.430,100)
}
```

Logarithmic transformation model

```
model {
for (i in 1:N) {
zeros[i] <- 0
phi[i] <- - loglike[i]
zeros[i] ~ dpois(phi[i])
A1[i]<- delta[i]*( beta1*sex[i]+ beta2*age[i] + beta3*tumorloco[i] +
beta4*tumorhemato [i] +  beta5*desnut[i] + beta6*comorbi[i] +
beta7*leucopenia[i])
A2[i]<- delta[i]*log(1+r*exp(beta1*sex[i]+ beta2*age[i] + beta3*tumorloco[i] +
beta4*tumorhemato [i] +  beta5*desnut[i] + beta6*comorbi[i] +
beta7*leucopenia[i])*Lambda[i])
A3[i]<- (log(1+r*exp(beta1*sex[i]+ beta2*age[i] + beta3*tumorloco[i] +
beta4*tumorhemato [i] +  beta5*desnut[i] + beta6*comorbi[i] +
```

```
beta7*leucopenia[i])*Lambda[i]))/r
loglike[i]<- A1[i] + delta[i]*log(lambda[i]) - A2[i] - A3[i]
lambda[i] <-a[i]
Lambda[i] <- a[i]*t[i]
a[i]~ dunif(0,0.1)
}
beta1<- log(theta1)
theta1~ dgamma(110.424,100)
beta2<- log(theta2)
theta2~ dgamma(101.315,100)
beta3<- log(theta3)
theta3~ dgamma(47.910,100)
beta4<- log(theta4)
theta4~ dgamma(99.274,100)
beta5<- log(theta5)
theta5~ dgamma(173.397,100)
beta6<- log(theta6)
theta6~ dgamma(168.632,100)
beta7<- log(theta7)
theta7~ dgamma(239.430,100)
r ~ dgamma(1,1)
}

list(t=c(162,10,2,182,182,12,52,10,3,4,9,7,6,182,182,182,182,90,182,20,10,
2,182,4,
28,46,5,3,182,12,4,31,182,182,4,2,3,182,182,182,182,182,3,2,73,43,43,4,45,
182,48,182,182,9,8,2,16,6,38,182,182,75,79,182,182,182,98,13,12,18,11,3,
182,53,2,182,110,182,9,38,29,182,100,68,5,7,19,182,45,182,182,2,44,182,8,
3,12,182,182,182,4,3,182,16,67,182,182,14,182,6,182,1,20,5,3,182,182,48,
182,21,12,142,182,182,182,10,182,13,10,20,12,3,182,23,182,20,8,57,6,2,86,
4,13,130,182,182,182,12,160,182,8,17,3,6,2,14,2,38,182,9,182,7,47,182,16,
182,1,58,52,17,159,182,10,4,28,100,182,17,182,182,4,182,182,2,26,182,182,
27,17,2,7,128,4,9,7,34,56,15,182,45,2,16,115,182,182,182,7,182,14,3,19,13,
4,182,164,3,6,2,105,3,7,182,55,74,182,182,182,6,182,37,182,182,53,182,182,
182,15,18,182,49,2,108,22,14,182,1,9,182,182,182,1,182,182,182,9,182,39,17
,15,97,182,182,182,182,182,182,182,182,2,182,14,4,19,17,3,182,182,173,182,
6,64,182,2,39,182,24,18,3,182,182,182,1,173,182,86,182,7,7,20,182,5,182,10
,14,12,2,182,3,3,4,28,11,149,182,11,25,182,182,182,7,182,182,182,147,8,6,
21,182,23,182,182,2,182,182,182,2,182,182,17,24,38,21,182,6,182,182,182,
67,182,3,7,182,8,182,182,23,182,182,182,182,182,5,5,34,7,25,3,6,170,25,
182,17,182,182,1,63,4,182,13,182,182,3,182,182,182,4,1,29,182,182,13,3,
61,84,8,182,64,93,182,126,182,182,21,10,182,10,3,182,7,182,24,182,37,34,
2,2,182,182,182,17,182,182,182,182,1,4,182,182,4,182,182,2,182,182,182,
182,15,182,14,182,26,182,182,48,18,182,182,38,182,48,144,134,182,28,182,
182,1,182,12,10,3,182,6,2,182,182,182,24,182,182,9,182,125,42,182,26,182,
148,5,15,182,182,18,182,6,2,182,182,3,70,61,8,1,182,182,182,182,18,182,5,
182,182,5,182,10,19,28,41,10,13,182,119,182,182,182,182,55,182,182,182,2,
3,182,7,4,3,50,14,22,182,182,2,182,182,8,182,1,8,182,182,182,182,56,1,123
,51,182,182,182,182,4,3,4,182,182,182,182,1,10,182,182,182,182,7,182,182,
3,2,46,108,182,182,97,113,182,182,182,155,155,101,7,6,182,182,182,22,59,1
```

```
8,94,12,27,182,5,182,182,182,182,9,14,182,182,58,2,2,1,8,182,6,182,182,35
,2,181,182,163,16,182,2,182,9,5,3,5,12,2,2,12,182,33,182,182,4,5,141,81,3
,178,182,182,11,182,25,27,27,182,49,182,3,2,19,182,16,182,19,6,5,3,20,182
,182,22,182,182,182,182,14,182,2,14,32,182,1,8,182,182,9,182,12,6,182,182
,182,182,182,31,182,182,26,182,78,3,10,182,31,1,16,6,14,182,99,182,182,18
2,182,100,182,182,30,182,182,89,72,7,182,58,182,182,42,182,8,5,29,182,58,
182,182,182,6,2,9,182,11,22,182,182,182,182,3,159,182,22,54,23,8,3,8,3,5,
182,182,2,1,25,40,182,1,182,44,182,182,182,182,182,31,12,182,15,182,182,5
,8,25,40,182,182,182,5,46,6,1,4,2,1,3,6,3,4,5,97,44,182,182,182,40,6,182,
173,182,182,36,182,86,99,66,182,33,1,13,5,7,1,5,11,182,9,182,2,182,182,50
,182,182,182,182,182,9,13,1,182,182,14,182,11,2,182,67,182,182,182,13,182
,182,29,182,182,4,10,182,5,36,43,38,22,15,66,182,182,182,182,182,15,182),
sex=c(0,1,1,1,1,1,0,1,1,1,1,0,0,1,1,1,1,1,1,1,1,1,1,0,1,1,0,1,0,1,1,0,0
,0,0,0,0,0,1,1,1,0,0,1,0,1,0,0,0,1,0,1,0,0,1,1,1,1,0,0,1,1,1,1,0,1,1,0,1,
1,1,1,1,0,0,1,0,0,1,1,0,0,1,0,1,0,0,1,1,1,0,1,1,0,0,0,0,0,1,1,1,1,1,1,1,1
,1,1,1,1,1,1,1,1,1,1,0,0,1,0,0,0,0,0,0,1,1,1,0,1,1,1,1,1,1,0,1
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,1,1,0,0,0,1,
1,0,0,1,0,1,0,0,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1,0,1,1,1,1,1,0,1,0,0,1,1
,1,1,0,0,0,1,1,1,0,1,0,1,0,0,0,0,1,0,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,
1,0,1,1,0,1,0,0,1,0,1,0,0,1,1,0,1,0,0,0,0,1,1,1,1,1,1,0,0,1,1,0,1,1,1,1,1,0,1
,1,0,1,1,1,1,1,1,1,1,0,0,0,1,1,0,0,0,0,1,0,1,1,0,0,0,1,1,1,1,1,0,1,0,1,0,0,
0,0,1,1,1,1,1,0,0,1,1,0,0,0,1,1,0,1,0,0,1,0,1,0,1,0,0,0,1,0,0,0,1,1,1,1,1,0
,0,0,1,0,1,0,0,1,1,0,1,0,1,0,1,1,0,0,0,1,0,1,1,1,1,0,1,0,1,1,1,0,1,1,0,1,
1,1,0,0,0,0,1,1,1,0,0,0,1,0,1,1,0,0,0,0,0,0,0,1,1,0,0,0,1,1,1,0,1,0,0,1,1,1
,1,0,0,0,0,0,0,1,1,1,1,1,1,0,1,1,0,0,1,0,0,0,0,1,1,1,1,1,1,1,1,0,1,1,1,1,1,
0,1,1,0,0,0,1,1,1,0,1,1,1,0,1,1,0,1,0,0,0,0,0,1,0,0,1,0,1,0,1,1,0,0,0,0,0,0
,0,1,1,1,0,0,1,1,0,0,1,1,1,0,1,0,0,0,1,1,1,1,1,1,0,0,0,1,1,0,0,1,0,0,0,1,1,
0,1,0,0,0,1,0,0,0,1,0,1,0,1,1,1,1,1,1,0,0,1,0,1,0,1,0,0,0,1,0,1,1,1,0,1,0,1
,1,0,1,1,0,1,0,0,0,1,0,1,0,1,0,0,0,0,0,1,0,1,1,0,1,0,1,1,1,1,1,0,1,1,0,1,0,1,
1,0,1,1,1,1,1,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,0,0,1,1,0,0,1,1,0,0,1,0,0,0,1,1,0
,1,0,1,1,0,0,1,1,1,0,0,0,0,1,1,0,0,1,1,1,1,0,0,1,0,0,1,0,1,1,0,1,1,0,1,0,1,
1,1,0,1,0,1,1,0,1,1,1,1,0,1,0,0,1,1,1,1,1,1,0,1,1,0,0,0,1,0,1,0,0,0,1,1,0,1,1
,0,0,1,1,0,0,1,1,0,0,1,0,0,1,0,1,1,0,1,1,0,1,0,1,0,1,1,1,0,0,0,1,1,0,1,1,
1,0,0,0,1,1,1,1,1,1,1,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,0,1,1,1,0,1,0,0,0,0,1,1,1,0
,1,0,0,0,1,0,1,1,1,1,1,0,0,0,1,1,0,1,0,0,0,1,1,1,0,1,0,1,1,0,0,1,1,1,1,1,1,
1,1,1,1,0,0,0,0,0,1,0,0,0,0,0,1,1,1,0,1,1,0,1,0,0,0,1,1,1,1,1),
delta=c(1,1,1,0,0,1,1,1,1,1,1,1,1,0,0,0,0,0,1,0,1,1,1,0,1,1,1,1,1,1,0,1,1,1,0
,0,1,1,1,0,0,0,0,0,1,1,1,1,1,1,1,1,0,1,0,0,1,1,1,1,1,1,1,0,0,1,1,0,0,0,1,1,1,
1,1,1,0,1,1,0,1,0,1,1,1,1,0,1,1,1,1,1,1,0,1,0,0,1,1,0,1,1,1,1,0,0,0,1,1,0,1,1,0
,0,1,0,1,0,1,1,1,1,1,1,0,0,1,0,1,1,1,1,0,0,0,1,0,1,1,1,1,1,1,0,1,0,1,1,1,1,1,1,1,
1,1,1,0,0,0,1,1,0,1,1,1,1,1,1,1,1,1,1,0,1,0,1,1,0,1,0,1,1,1,1,1,1,0,1,1,1,1,1,0,1,0
,0,1,0,0,1,1,0,0,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,0,0,0,0,1,0,1,1,1,1,1,1,0,
1,1,1,1,1,1,1,1,0,1,1,0,0,0,1,0,1,0,0,1,0,1,0,0,0,1,1,0,1,1,1,1,1,1,0,1,1,0,0,0,
1,0,0,0,1,0,1,1,1,1,1,0,0,0,0,0,0,0,0,0,1,0,1,1,1,1,1,1,0,0,1,0,1,1,0,1,1,0,1,
1,1,0,0,0,1,1,0,1,0,1,1,1,1,0,1,0,1,1,1,1,1,1,1,1,0,1,1,0,1,1,0,0,0,1,0,0,
0,1,1,1,1,1,0,1,0,0,1,0,0,0,1,0,0,1,1,1,1,1,0,1,0,0,0,1,0,1,1,0,1,0,0,1,0,0,
0,0,0,1,1,1,1,1,1,1,1,1,1,1,0,1,0,0,1,1,1,0,1,0,0,1,0,0,0,1,1,1,1,0,0,1,1,1,1,
1,0,1,1,0,1,0,0,1,1,0,1,1,0,1,0,1,0,1,1,1,1,1,0,0,0,1,0,0,0,0,0,1,1,0,0,1,0,
0,1,0,0,0,0,1,0,1,0,1,0,0,1,1,0,0,1,0,1,1,1,1,0,1,0,0,1,0,1,1,1,0,1,1,0,0
```

```
,0,1,0,0,1,0,1,1,0,1,0,1,1,1,0,0,1,0,1,1,0,0,1,1,1,1,1,0,0,0,0,1,0,1,0,0,
1,0,1,1,1,1,1,1,0,1,0,0,0,0,1,0,0,0,1,1,0,1,1,1,1,1,1,1,0,0,1,0,0,1,0,1,1,
0,0,0,0,1,1,1,1,0,0,0,0,1,1,1,0,0,0,0,1,1,0,0,0,0,1,0,0,1,1,1,1,1,0,0,1,1,
0,0,0,1,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,0,1,0,0,0,0,0,1,1,0,0,1,1,1,1,1,1,0,
1,1,1,0,1,1,0,1,0,1,1,1,1,1,1,1,1,1,0,1,0,0,1,1,1,1,1,1,1,0,0,1,0,1,1,1,1,
0,1,1,1,0,1,0,1,1,1,1,1,1,0,0,1,0,0,0,0,1,0,1,1,1,0,1,1,0,0,1,0,1,1,1,0,0,
0,0,1,0,0,1,0,1,1,1,1,0,1,1,1,1,1,1,0,1,0,0,0,0,1,0,0,1,0,0,1,1,1,1,0,1,0,
0,1,1,1,0,1,0,0,0,1,1,1,0,1,1,0,0,0,0,1,1,0,1,1,1,1,1,1,1,1,1,0,0,1,1,1,1,
0,1,0,1,0,0,0,0,0,1,1,0,1,0,0,1,1,1,1,1,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,
0,0,0,1,1,0,1,0,0,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,0,1,0,1,0,0,1,0,0,0,0,0,1,
1,1,0,0,1,0,1,1,0,1,0,0,0,1,0,0,1,0,0,1,1,0,1,1,1,1,1,1,1,1,0,0,0,0,0,1,0),
age=c(69,67,69,35,63,63,53,75,68,78,69,66,70,60,74,64,36,40,68,61,74,65,
55,67,73,64,71,46,68,74,52,60,63,57,43,18,30,64,45,29,67,22,47,50,64,45,
26,38,67,76,22,50,46,49,72,69,76,41,67,54,62,69,86,62,36,78,50,83,88,19,
62,25,71,67,77,45,21,59,71,77,50,55,50,60,74,70,35,87,48,54,60,22,85,69,
53,69,79,18,58,40,63,71,49,73,52,70,34,68,59,33,51,78,43,67,43,83,72,68,
67,69,70,45,58,57,41,63,40,58,53,54,51,83,73,73,58,30,41,27,69,76,62,47,
58,39,79,59,75,59,70,62,60,79,30,71,71,65,68,54,40,45,56,43,67,51,58,52,
58,58,69,60,61,73,67,52,57,69,54,74,70,73,87,54,43,71,77,64,62,63,48,54,
47,68,79,64,51,72,73,87,79,42,57,58,77,30,20,46,59,62,75,49,52,73,51,34,
21,50,81,80,66,51,41,18,59,22,48,50,57,19,57,23,78,58,63,70,77,55,45,65,
69,66,73,60,71,54,56,67,75,26,50,53,70,63,64,45,73,58,83,72,40,20,23,30,
67,49,67,55,30,61,44,70,69,50,45,64,34,20,63,25,46,51,75,24,67,48,61,45,
49,38,60,39,65,77,64,62,79,69,70,52,74,47,58,69,63,59,68,68,50,16,65,54,
66,73,43,38,55,79,45,74,36,25,43,62,53,57,72,44,81,79,56,68,52,19,28,66,
71,27,71,38,50,72,65,80,81,74,66,83,70,65,62,55,45,48,64,31,51,49,40,38,
46,72,62,50,52,53,53,44,74,71,21,72,71,65,77,75,60,66,68,50,70,30,57,73,
57,48,36,60,62,48,55,56,76,80,48,38,55,59,78,25,77,64,54,45,28,76,21,31,
78,27,69,67,77,87,43,32,46,69,70,52,42,73,27,29,80,65,60,47,20,74,64,78,
67,64,33,53,38,22,60,66,85,66,82,31,49,91,32,51,56,67,56,56,22,60,57,49,
32,57,38,60,57,74,80,53,30,22,54,76,69,42,55,19,61,20,85,64,59,19,41,70,
59,84,60,62,64,45,27,64,54,88,38,52,72,68,23,73,73,67,52,65,61,52,46,64,
55,50,83,58,54,45,29,20,30,52,70,53,75,63,68,71,52,72,54,44,29,67,77,19,
78,69,71,83,43,65,38,17,57,61,52,58,72,75,30,72,65,45,45,53,72,53,53,91,
64,85,68,56,17,63,60,82,70,39,73,73,63,48,58,47,63,85,65,73,71,68,84,38,
50,32,29,73,43,68,63,67,51,35,67,37,62,53,61,52,50,45,84,59,26,45,63,64,
66,54,34,67,47,65,82,55,52,85,61,60,85,62,64,71,53,73,57,55,54,62,77,73,
61,56,73,49,77,18,73,68,63,55,76,36,73,70,65,64,33,59,48,76,57,72,63,66,
83,75,65,89,28,70,70,54,69,69,63,28,54,50,58,60,49,65,64,90,40,60,74,46,
45,68,66,64,39,59,73,81,46,61,57,48,54,75,81,69,39,67,69,69,42,75,70,56,
52,38,37,61,62,76,53,40,51,59,73,76,59,67,21,73,64,71,51,24,69,23,18,69,
66,63,72,70,64,79,45,51,40,67,65,60,73,55,64,45,65,72,69,69,48,86,77,73,
60,63,49,61,23,51,18,41,64,63,69,62,59,34,63,65,60,49,40,65,74,74,77,53,
44,72,77,31,60,33,51,61,65,57,51,29,44,54,70,72,73,70,49,88,33,77,72,62,
68,57,70,57,98,55,52,73,36,64,72,41,69,24,53,53,50,74,44,19,79,64,54,63,
46,38,68,62,58,33,75,76,63,55,69,58,68,21,77,57,67,80,74,63,54,60,42,79,
70,41,46,53,48,72,66,71,59,67,27,70,66,61,66,75,61,38,63,74,68,70,59,66),
tumorloco=c(1,0,0,1,1,0,1,1,1,0,1,0,1,1,0,1,1,0,1,0,0,1,1,1,1,1,1,1,1,1,1,
0,1,1,1,0,0,0,1,1,1,0,1,0,0,1,1,0,1,0,0,1,1,1,0,1,1,1,0,1,1,1,0,1,1,0,0,
```

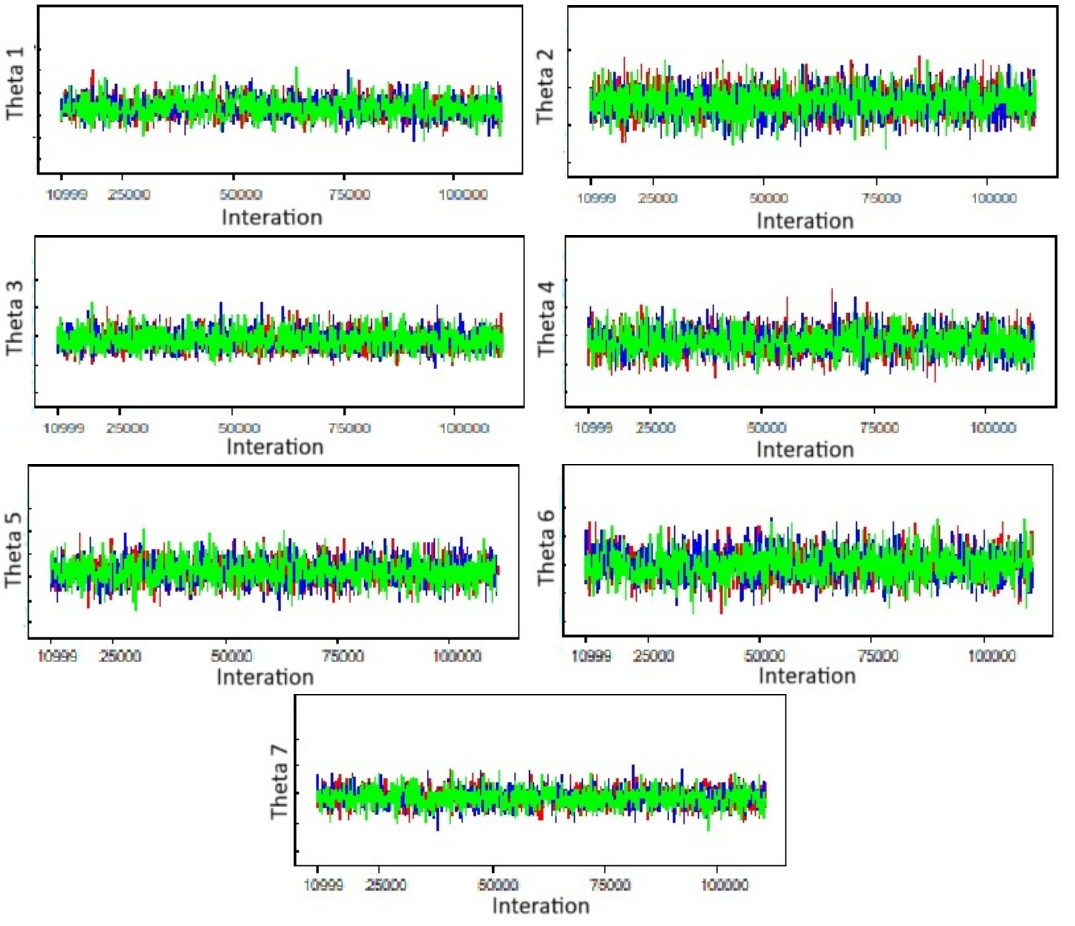
```
0,1,1,1,0,0,1,1,0,1,0,0,1,1,0,1,1,0,0,0,0,1,1,1,1,0,1,1,0,0,0,0,1,1,1,1,
1,0,1,1,1,0,1,1,1,1,0,0,0,1,1,1,1,1,0,0,0,1,1,0,0,0,0,1,0,0,1,1,1,0,0,0,1,
1,0,1,0,0,1,1,1,1,1,1,1,1,0,1,0,1,0,1,1,0,1,0,0,0,0,1,0,0,0,1,0,1,0,1,1,1,
0,1,0,0,1,0,1,0,1,1,1,1,1,1,0,1,1,0,1,0,0,1,1,1,1,1,1,0,0,0,1,1,0,1,1,1,0,
0,0,0,1,0,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1,0,1,0,0,0,0,0,0,1,1,0,0,0,1,0,
0,1,1,1,0,1,1,1,0,0,1,1,1,1,1,1,0,0,1,1,1,1,1,1,1,1,1,0,0,1,0,0,1,0,1,1,1,0,
0,1,1,1,1,0,1,1,1,1,0,1,1,1,0,1,1,1,0,1,1,1,1,1,1,1,1,0,0,1,1,1,1,0,1,1,1,
1,0,1,1,1,1,1,1,1,1,1,0,1,1,0,0,1,1,0,1,1,0,0,0,0,0,1,0,1,0,1,0,1,0,0,1,1,0,
1,1,1,0,0,1,1,0,1,0,0,1,0,1,1,1,1,1,0,0,1,1,0,1,1,0,0,1,1,1,1,1,1,1,0,1,1,
0,0,0,1,1,1,0,0,1,1,1,1,1,0,1,1,0,1,1,0,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1,1,1,0,
0,1,0,1,1,0,1,1,1,0,1,0,0,1,1,0,1,1,1,1,1,1,0,1,1,0,1,0,1,1,1,1,1,1,0,1,1,1,
0,0,0,1,1,0,1,1,0,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1,0,0,0,1,0,1,1,0,1,1,
1,0,1,1,1,1,0,1,1,0,1,1,1,1,1,1,1,1,1,1,1,0,1,0,0,1,1,1,0,1,0,1,1,1,0,0,1,
1,1,0,0,1,1,1,1,1,1,0,0,0,1,1,1,0,1,0,1,1,1,1,1,1,0,0,1,1,1,1,1,0,0,1,0,1,1,
1,1,1,0,0,0,0,0,1,0,1,1,1,1,1,1,1,1,0,1,0,0,0,0,0,1,1,1,1,1,0,1,1,1,0,1,0,1,1,0,
1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,0,0,0,0,0,0,0,1,1,1,1,1,0,1,1,0,0,0,1,1,1,
0,1,1,1,1,1,1,1,0,1,0,1,0,1,1,1,1,0,1,0,1,1,1,0,1,1,1,1,0,1,0,0,1,1,1,1,1,0,1,1,0,
0,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,0,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,0,1,
1,1,0,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,0,1,0,1,1,1,1,1,1,0,0,
0,0,1,0,0,0,1,1,1,1,1,1,1,1,0,0,1,0,1,1,1,1,1,0,0,0,1,1,1,1,1,0,0,1,1,0,0,0,1,
1,0,1,0,1,1,1,1,1,0,1,1,1,0,1,1,1,1,1,1,0,0,0,0,0,1,0,0,1,1,1,1,1,1,0,0,0,1,
1,1,1,0,1,0,1,1,1,1,1,0,1,1,1,1,1,0,1,1,1,0,0,0,1,1,0,1,0,0,0,0,0,0,0,1,1,1,1,
1,1,1,1,0),
tumorhemato=c(0,0,1,0,0,1,0,0,0,1,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,1,1,0,0,0,1,0,0,0,0,0,1,0,1,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,
1,1,0,0,0,0,1,0,0,1,0,1,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,0,0,
0,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,1,0,1,1,0,0,0,0,0,1,1,
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,1,1,0,0,0,0,0,0,1,0,0,0,1,0,0,
0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,1,1,0,0,1,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,1,0,0,1,0,0,0,0,0,
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,1,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,
1,1,0,1,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,
0,1,1,0,0,0,1,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,1,0,0,1,0,
0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,1,
0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,1,0,0,0,
0,0,0,0,1,1,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,1,1,1,0,0,0,0,1,0,0,0,0,0,1,0,0,
0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,
0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,
0,0,0,0,1,1,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,1,1,0,0,1,1,1,
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,1,0,
0,0,0,0,1,0,1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,1,1,1,1,0,1,0,0,0,
0,0,0,0,1),
desnut=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
```

```
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,
0,1,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,1,0,0,1,0,0,0,0,0,1,1,0,0,0,0,1,0,0,
0,1,0,1,0,0,1,0,0,0,1,1,0,1,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,1,0,1,0,1,0,1,1,1,0,0,1,1,
0,1,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,
1,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,
1,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,1,0,0,1,0,0,0,0,0,0,
1,0),
comorbi=c(0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,
1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,
0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,1,0,0,1,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,
0,0,0,0,1,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
1,0,0,0,0,1,0,1,0,0,0,0,0,1,0),
leucopenia=c(0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0
```

```
,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1
,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0
,0,1,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0,0,1,0,0,0,0,0,0
,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0
,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0
,1,1,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0
,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0
,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),N = 862)
```

Appendix 2 Convergence of the MCMC (model 3)

Trace plots (model 3)



GR index (model 3)