Afrika Statistika

Afrika Statistika
Vol. 15 (1), 2020, pages 2199 - 2224.
DOI: http://dx.doi.org/10.16929/as/2020.2199.154

ISSN 2316-090X

Gamma and Inverse Gaussian Frailty Models with Time-varying co-variates Based on Some Parametric Baseline Hazards Functions

Abiodun Waliyu Oyekunle [1*], Kazeem Adedayo Adeleke and Akinlolu Adeseye Olosunde [3]

[1,2,3]Department of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria.

Received October 9, 2019; Accepted December 18, 2019

Abstract. Ignoring the existence of frailty term in the analysis of survival time data, when heterogeneity is present will produce a less accurate estimated parameters with higher standard errors. In survival analysis, Cox proportional hazards model is frequently used to measure the effects of covariates. The covariates may fail to fully account for the true differences in hazard. This may be due to an existence of another response variable that is disregarded in the model but can be explained by the term known as frailty. The incorporation of frailty in the model thereby avoid underestimation and overestimation of parameters and also correctly measure the effects of the covariates on the response variable. This paper presents a parametric non-proportional hazard models with Weibull, Loglogistic and Gompertz as baseline distributions and Gamma and Inverse Gaussian as frailty distribution. A maximum likelihood method is used and is illustrated with a numerical example in which the fit is compared using Akaike Information Criterion (AIC):

Key words: time-varying co-variate; unobserved co-variates; non-proportional hazard model; inverse Cumulative Hazard; survival time; frailty models; Akaike Information Criterion.
AMS 2010 Mathematics Subject Classification Objects : 62H2; 62F10

---

*Corresponding author: Oyekunle, A. W (abiodunoyekunle5@gmail.com)
Adeleke, K. A : adedayobright@gmail.com
Olosunde, A. A : akinolosunde@gmail.com

Résumé. (French) Ignorer l'existence d'un terme de fragilité dans l'analyse des données de temps de survie, lorsque l'hétérogénéité est présente, peut aboutir a des estimateurs de paramètres moins précis avec des erreurs standard élevées. Dans l'analyse de survie, le modèle des risques proportionnels de Cox est fréquemment utilisé pour mesurer les effets des covariables. Les covariables peuvent ne pas tenir pleinement compte des vraies différences de risque. Cela peut être dû à l'existence d'une autre variable de réponse qui n'est pas prise en compte dans le modèle mais peut être expliquée par un terme connu sous le labeli fragile. L'incorporation de la fragilité dans le modèle évite ainsi la sous-estimation et la surestimation des paramètres et mesure également correctement les effets des covariables sur la variable de réponse. Cet article présente des modèles de risques paramétriques non proportionnels avec Weibull, Loglogistic et Gompertz comme distributions de base et Gamma et Gaussian inverse comme distribution de fragilité. La méthode du maximum de vraisemblance est utilisée et la procédure est illustrée par un exemple numérique et l'ajustement est comparé à l'aide du critère d'information Akaike (AIC).

The authors.

(1) Abiodun Waliyu Oyekunle, M.Sc., is a Researcher at Obafemi Awolowo University, Ile-Ife, Nigeria, (Department of Mathematics).

(2) Kazeem Adedayo Adeleke, (Ph.D), is a Lecturer at Obafemi Awolowo University, Ile-Ife, Nigeria, (Department of Mathematics).

(3) Akinlolu Adeseye Olosunde, (Ph.D), is a Senior Lecturer at Obafemi Awolowo University, Ile-Ife, Nigeria, (Department of Mathematics).

1. Introduction

One of the advantages of the Cox proportional hazard model is to assess the co-variate effects on hazard function. However, the co-variate do not always account fully for the true differences in risk. Most importantly, in clustered survival time data. Hence, incorporating the frailty term into the Cox model enhances correct measurement of co-variate effect and avoid underestimation or overestimation of the parameters with standard errors in the analysis as stated found in Duchateau and Janssen (2008), Keyfitz and Littman (1979) and Lancaster (1979). The assumption of Proportional Hazard Model states that the hazard ratio is constant over time. However, in real life problem, the assumption is not always satisfied and that is why the hazard ratio varies with time (See Cox (1975)). In view of this, a stratified Cox and extended Cox regression models that use time-dependent and time-fixed co-variates are adopted, as discussed in Austin (2012), Kronborg and Aaby (1990), Zhang *et al.* (2018) and Zhou (2001).

Authors in Austin (2012) and Robert and Casella (2009) discussed a piecewise comparison of survival function in stratified proportional hazard model and presented a practical technique to generate random number from standard and non-standard distribution for non-proportional hazard model. Recently, a number of researches have been directed towards modeling time-varying co-variates as well as stratification which are semi-parametric non-proportional hazard models (Lehr (2004), Abrahamowicz *e*t al. (2007), Bender *e*t al. (2005), Austin (2012), Ata and Sozer (2007), Zhou (2001)). A more advanced method of generating time varying co-variate is the work of Zhou (2001) where the use of an exponential distribution was examined in conjunction with a transformation to the Cox model including time-varying co-variate. A piecewise exponential distribution was used to obtain a dichotomous or step function co-variate which was in turn incorporated into the Cox model and analyzed through a semi-parametric approach. Time-varying co-variate is very crucial in Cox model because it guides against the problem of survivor treatment.

The usefulness of Cox model with time-varying co-variates are likely to continue to become progressively vital in medical research. The methods put forth by Sylvester and Abrahamowicz (2007) are however not presented in a close form. Leemis (1987), Leemis *et al.* (1990), and Shih and Leemis (1993). have offered diverse frameworks for generation of survival time that follow a Cox model with time-varying following accelerated life and proportional hazards models where his procedures adopted one time-varying co-variate and no time fixed co-variates. A recent study on Cox regression model in the presence of non-proportional hazards was carried out by Ata and Sozer (2007), where they worked on alternative different models in the violation of proportional assumption.

Frailty models are random effect models for time-to-event data, where the random effect has a multiplicative effect on the baseline hazard function, Keyfitz and Littman (1979). It is an extension of the most popular Cox proportional hazard model. Absence of frailty term in Lancaster (1979) leads to underestimation of co-variates in the study of unemployed rates. Also, the work of Hougaard (2000) showed that turning a blind eye to frailty factor leads to regression coefficient estimates attenuation towards zero. Thus, the choice of frailty distribution is a very crucial aspect in the area of frailty models. Example of such distributions is one parameter gamma distribution Hougaard (1986). (see an interesting application of frailty models in in Fagbamigbe *et al.* (2019), for example).

The novel of this work from Adeleke *e*t al. (2015), Austin (2012), Zhang *et al.* (2018) lies in incorporation of time-varying co-variate and frailty term into the Cox model. The main aim of this paper is to estimate the frailty variance θ which is used to examine the degree of heterogeneity in the study population. This paper consider the estimation of time-dependent co-variate for Gamma and Inverse Gaussian frailty models for a right censored data. The assessment of our approach is based on the application of a real-life data set and the fit is compared using Akaike Information Criterion.

2. Dichotomous time-varying co-variates

The time at which time-varying co-variate changes from untreated to treated is denoted as $(y = 0)$ and $(y = 1)$ respectively where $y(t) = 0$ for $t < t_0$ and $y(t) = 1$ for $t \geq t_0$, where t_0 is the time at which the time-varying co-variates changed. The cumulative hazard function, inverse cumulative hazard function and survival time for these two domains with a frailty term z are derived below. The risk function is expressed as

$$m(t \mid x(t), z) = m_0(t)exp(\alpha' x + \beta_t' y(t) + k)$$

where $k = log(z)$. The cumulative hazard function is expressed as

$$M(t \mid x(t), z) = \int_0^t exp(\alpha' x + \beta_t' y(u) + k) m_0(u)\, du.$$

Proposition 1. *Let t be a random variable which has a Weibull distribution having scale parameter λ and a shape parameter γ for a dichotomous case of $t < t_0$, with hazard model $m(t \mid x(t), z) = \lambda\gamma t^{\gamma-1} \exp(\alpha' x + k)$. Then the cumulative hazard function is $M(t \mid x(t), z) = \lambda exp(\alpha' x + k)t^{\gamma}$ with inverse cumulative hazard $M^{-1}(t \mid x(t), z) = \left(\frac{t}{\lambda exp(\alpha' x+k)}\right)^{\frac{1}{\gamma}}$, if $t < \lambda exp(\alpha' x + k)t_0^{\gamma}$.*

Proof of Proposition 1. We have

$$\begin{aligned} M(t \mid x(t), z) &= \int_0^t exp(\alpha' x + k)\lambda\gamma u^{\gamma-1} du \\ &= \lambda\gamma exp(\alpha' x + k) \int_0^t u^{\gamma-1} du \\ &= \lambda\gamma exp(\alpha' x + k) \left[\frac{1}{\gamma} u^{\gamma}\right]_0^t \\ &= \lambda exp(\alpha' x + k) \left[u^{\gamma}\right]_0^t \end{aligned}$$

$$M(t \mid x(t), z) = \lambda exp(\alpha' x + k)t^{\gamma}. \quad (1)$$

Also, the inverse cumulative hazard function is:

$$M(t \mid x(t), z) = \lambda exp(\alpha' x + k)t^{\gamma}.$$

$$t^{\gamma} = \frac{M(t \mid x(t), z)}{\lambda exp(\alpha' x + k)}.$$

$$t = \left(\frac{M(t \mid x(t), z)}{\lambda exp(\alpha' x + k)} \right)^{\frac{1}{\gamma}}.$$

Thus,

$$M^{-1}(t \mid x(t), z) = \left(\frac{t}{\lambda exp(\alpha' x + k)} \right)^{\frac{1}{\gamma}}, \quad \text{if} \quad \mathrm{t} < \lambda \mathrm{exp}(\alpha' \mathrm{x} + \mathrm{k}) \mathrm{t}_0^{\gamma}. \tag{2}$$

Equations (1) and (2) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t < t_0$.

Proposition 2. *Let t be a random variable which has a Weibull distribution having scale parameter λ and a shape parameter γ for a dichotomous case of $t \geq t_0$, with hazard model*
$m(t \mid x(t), z) = \lambda \gamma t^{\gamma-1} \exp(\alpha' x + \beta_t' y(t) + k)$*. Then the cumulative hazard function is* $M(t \mid x(t), z) = \lambda exp(\alpha' x + k) \left[t_0^{\gamma} + \exp(\beta_t') t^{\gamma} - \exp(\beta_t') t_0^{\gamma} \right]$*, with inverse cumulative hazard* $M^{-1}(t \mid x(t), z) = \left(\frac{t - \lambda exp(\alpha' x + k) t_0^{\gamma} + \lambda exp(\alpha' x + k + \beta_t') t_0^{\gamma}}{\lambda \exp(\alpha' x + k + \beta_t')} \right)^{\frac{1}{\gamma}}$*, when* $t \geq \lambda \exp(\alpha' x + k) t_0^{\gamma}$.

Proof of Proposition 2. We have

$$\begin{aligned} M(t \mid x(t), z) &= \int_0^t exp(\alpha' x + \beta_t' y(u) + k) \lambda \gamma u^{\gamma-1} du \\ &= \lambda \gamma exp(\alpha' x + k) \int_0^t exp(\beta_t' y(u)) u^{\gamma-1} du \\ &= \lambda \gamma exp(\alpha' x + k) \left[\int_0^{t_0} exp(\beta_t' y(u)) u^{\gamma-1} du + \int_{t_0}^t exp(\beta_t' y(u)) u^{\gamma-1} du \right] \end{aligned}$$

Note that $y(u) = 0$ at interval $(0, t_0)$ and $y(u) = 1$ at interval (t_0, t)

$$M(t \mid x(t), z) = \lambda \gamma exp(\alpha' x + k) \left\{ \left[\frac{1}{\gamma} u^{\gamma} \right]_0^{t_0} + \exp(\beta_t') \left[\frac{1}{\gamma} u^{\gamma} \right]_{t_0}^t \right\}$$

$$M(t \mid x(t), z) = \lambda exp(\alpha' x + k) \left[t_0^{\gamma} + \exp(\beta_t') t^{\gamma} - \exp(\beta_t') t_0^{\gamma} \right]. \tag{3}$$

Also, the inverse of the cumulative hazard function in the domain $M(t \mid x(t), z) \geq \lambda exp(\alpha' x + k) t_0^{\gamma}$ is:

$$M(t \mid x(t), z) - \lambda exp(\alpha' x + k) t_0^{\gamma} + \lambda exp(\alpha' x + k) \exp(\beta_t') t_0^{\gamma} = \lambda \exp(\alpha' x + k) \exp(\beta_t') t^{\gamma}$$

$$t = \left(\frac{M(t \mid x(t), z) - \lambda exp(\alpha' x + k) t_0^{\gamma} + \lambda exp(\alpha' x + k + \beta_t') t_0^{\gamma}}{\lambda \exp(\alpha' x + k + \beta_t')} \right)^{\frac{1}{\gamma}}$$

$$M^{-1}(t \mid x(t), z) = \left(\frac{t - \lambda exp(\alpha' x + k) t_0^{\gamma} + \lambda exp(\alpha' x + k + \beta_t') t_0{}^{\gamma}}{\lambda \exp(\alpha' x + k + \beta_t')} \right)^{\frac{1}{\gamma}}, \qquad (4)$$

when $t \geq \lambda \exp(\alpha' x + k) t_0^{\gamma}$. Equations (3) and (4) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t \geq t_0$. Hence, the Cox-Weibull survival time is

$$\mathrm{T} = \begin{cases} \left(\frac{-\log(\mathrm{u})}{\lambda \exp(\alpha' \mathrm{x} + \mathrm{k})} \right)^{\frac{1}{\gamma}}, & \text{if} \quad -\log(\mathrm{u}) < \lambda \exp(\alpha' \mathrm{x} + \mathrm{k}) \mathrm{t}_0^{\gamma} \\ \left(\frac{-\log(\mathrm{u}) - \lambda \exp(\alpha' \mathrm{x} + \mathrm{k}) \mathrm{t}_0^{\gamma} + \lambda \exp(\alpha' \mathrm{x} + \mathrm{k} + \beta_{\mathrm{t}}') \mathrm{t}_0^{\gamma}}{\lambda \exp(\alpha' \mathrm{x} + \mathrm{k} + \beta_{\mathrm{t}}')} \right)^{\frac{1}{\gamma}}, & \text{if} \quad -\log(\mathrm{u}) \geq \lambda \exp(\alpha' \mathrm{x} + \mathrm{k}) \mathrm{t}_0^{\gamma} \end{cases} \qquad (5)$$

where $u \sim U(0, 1)$.

Proposition 3. *Let t be a random variable which follows a Log-logistic distribution having a scale parameter λ and a shape parameter γ for a dichotomous case of $t < t_0$, with hazard model $m(t \mid x(t), z) = \frac{\exp(\gamma) \lambda t^{\lambda - 1}}{1 + \exp(\gamma) t^{\lambda}} \exp(\alpha' x + k)$. Then the cumulative hazard function is $M(t \mid x(t), z) = \exp(\alpha' x + k) \ln(1 + \exp(\gamma) t^{\lambda})$, with inverse cumulative hazard as*

$$M^{-1}(t \mid x(t), z) = \left(\frac{1}{\exp(\gamma)} \exp\left(\frac{t}{\exp(\alpha' x + k)} \right) - 1 \right)^{\frac{1}{\lambda}}.$$

Proof of Proposition 3. We have

$$M(t \mid x(t), z) = \int_0^t \frac{\exp(\gamma) \lambda u^{\lambda - 1}}{1 + \exp(\gamma) u^{\lambda}} \exp(\alpha' x + k) du$$

$$M(t, x, y(t)) = \lambda \exp(\alpha' x + k) \exp(\gamma) \int_0^t \frac{u^{\lambda - 1}}{1 + \exp(\gamma) u^{\lambda}} du$$

Consider the integral part

$$\int_0^t \frac{u^{\lambda - 1}}{1 + \exp(\gamma) u^{\lambda}} du$$

Let $p = 1 + \exp(\gamma u^{\lambda})$

$$du = \frac{dp}{\lambda \exp(\gamma) u^{\lambda-1}}$$

$$\int_0^t \frac{u^{\lambda-1} dp}{p\lambda \exp(\gamma) u^{\lambda-1}}$$

$$= \frac{1}{\lambda \exp(\gamma)} \ln p$$

$$= \left[\frac{\ln(1+\exp(\gamma)u^{\lambda})\exp(-\gamma)}{\lambda}\right]_0^t$$

$$M(t \mid x(t), z) = \lambda exp(\alpha' x + k)\exp(\gamma)\left[\frac{\ln(1+\exp(\gamma)u^{\lambda})\exp(-\gamma)}{\lambda}\right]_0^t$$

$$M(t \mid x(t), z) = \exp(\alpha' x + k)\ln(1+\exp(\gamma)t^{\lambda}) \tag{6}$$

Also, the inverse of the cumulative hazard function when

$$M(t \mid x(t), z) < \exp(\alpha' x + k)\ln(1+\exp(\gamma)t^{\lambda})$$

is:

$$M(t \mid x(t), z) = \exp(\alpha' x + k)\ln(1+\exp(\gamma)t^{\lambda})$$

$$\frac{M(t \mid x(t), z)}{\exp(\alpha' x + k)} = \ln(1+\exp(\gamma)t^{\lambda})$$

$$\exp\left(\frac{M(t \mid x(t), z)}{\exp(\alpha' x + k)}\right) - 1 = \exp(\gamma)t^{\lambda}$$

$$t = \left(\frac{1}{\exp(\gamma)}\exp\left(\frac{M(t \mid x(t), z)}{\exp(\alpha' x + k)}\right) - 1\right)^{\frac{1}{\lambda}}$$

$$M^{-1}(t \mid x(t), z) = \left(\frac{1}{\exp(\gamma)}\exp\left(\frac{t}{\exp(\alpha' x + k)}\right) - 1\right)^{\frac{1}{\lambda}}. \tag{7}$$

Equations (6) and (7) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t < t_0$.

Proposition 4. *Let t be a random variable which follows a log-logistic distribution having a scale parameter λ and a shape parameter γ for a dichotomous case of $t \geq t_0$, with hazard model $m(t \mid x(t), z) = \frac{\exp(\gamma)\lambda t^{\lambda-1}}{1+\exp(\gamma)t^{\lambda}} \exp(\alpha' x + \beta'_t y(t) + k)$. Then the cumulative hazard function is* $M(t \mid x(t), z) = exp(\alpha' x + k)\Big[\ln(1+\exp(\gamma)t_0^{\lambda}) + \exp(\beta'_t)\ln(1+\exp(\gamma)t^{\lambda}) - \exp(\beta'_t)\ln(1+\exp(\gamma)t_0^{\lambda})\Big]$ *with inverse cumulative hazard*

$$
\begin{aligned}
M^{-1}(t \mid x(t), z) \;=\; & \left(\frac{1}{\exp(\gamma)}\left[\left[\exp\left(\frac{t - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta'_t)\exp(\alpha' x + k)}\right)\right.\right.\right. \\
& \times \left.\left.\left. \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta'_t)\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta'_t)\exp(\alpha' x + k)}\right)\right] - 1\right]\right)^{1/\lambda}.
\end{aligned}
$$

Proof of Proposition 4. We have

$$
\begin{aligned}
M(t \mid x(t), z) &= \int_0^t \frac{\exp(\gamma)\lambda u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} \exp(\alpha' x + \beta'_t y(u))du \\
&= \lambda \exp(\alpha' x + k)\exp(\gamma) \int_0^t \frac{u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} \exp(\beta'_t y(u))du \\
&= \lambda \exp(\alpha' x + k)\exp(\gamma) \Big[\int_0^{t_0} \frac{u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} \exp(\beta'_t y(u))du \\
&\quad + \int_{t_0}^t \frac{u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} \exp(\beta'_t y(u))du \Big] \\
&= \lambda \exp(\alpha' x + k)\exp(\gamma) \Big[\int_0^{t_0} \frac{u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} du \\
&\quad + \exp(\beta'_t) \int_{t_0}^t \frac{u^{\lambda-1}}{1+\exp(\gamma)u^{\lambda}} du \Big] \\
&= \lambda \exp(\alpha' x + k)\exp(\gamma) \Bigg\{ \left[\frac{\ln(1+\exp(\gamma)u^{\lambda})\exp(-\gamma)}{\lambda}\right]_0^t \\
&\quad + \exp(\beta'_t)\left[\frac{\ln(1+\exp(\gamma)u^{\lambda})\exp(-\gamma)}{\lambda}\right]_{t_0}^t \Bigg\}
\end{aligned}
$$

$$
M(t \mid x(t), z) = exp(\alpha' x + k)\Big[\ln(1+\exp(\gamma)t_0^{\lambda}) + \exp(\beta'_t)\ln(1+\exp(\gamma)t^{\lambda}) - \exp(\beta'_t)\ln(1+\exp(\gamma)t_0^{\lambda})\Big] \quad (8)
$$

Also, the inverse of the cumulative hazard function when $M(t \mid x(t), z) \geq \exp(\alpha' x + k)\ln(1+\exp(\gamma)t^{\lambda})$ is:

$$M(t \mid x(t), z) - exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda}) + exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda}) = \exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})$$

$$\ln(1+\exp(\gamma)t^{\lambda}) = \frac{M(t \mid x(t), z) - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)} + \frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}$$

Taking the exponent of both sides leads to

$$1+\exp(\gamma)t^{\lambda} = \exp\left(\frac{M(t \mid x(t), z) - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right) \times \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right)$$

$$\exp(\gamma)t^{\lambda} = \left[\exp\left(\frac{M(t \mid x(t), z) - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right) \times \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right)\right] - 1$$

$$t^{\lambda} = \frac{1}{\exp(\gamma)}\left[\left[\exp\left(\frac{M(t \mid x(t), z) - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right) \times \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right)\right] - 1\right]$$

$$t = \left(\frac{1}{\exp(\gamma)}\left[\left[\exp\left(\frac{M(t \mid x(t), z) - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right) \times \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right)\right] - 1\right]\right)^{\frac{1}{\lambda}}$$

$$M^{-1}(t \mid x(t), z) = \left(\frac{1}{\exp(\gamma)}\left[\left[\exp\left(\frac{t - \exp(\alpha' x + k)\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right) \times \exp\left(\frac{\exp(\alpha' x + k)\exp(\beta_t')\ln(1+\exp(\gamma)t^{\lambda})}{\exp(\beta_t')\exp(\alpha' x + k)}\right)\right] - 1\right]\right)^{\frac{1}{\lambda}} \tag{9}$$

Equations (8) and (9) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t < t_0$.

Hence, the Cox-Log-logistic survival time is:

$$\mathrm{T} = \begin{cases} \left(\frac{1}{\exp(\gamma)} \exp\left(\frac{-\log(\mathrm{u})}{\exp(\alpha'\mathrm{x}+\mathrm{k})}\right) - 1\right)^{\frac{1}{\lambda}}, & \text{if} \quad -\log(\mathrm{u}) < \exp(\alpha'\mathrm{x}+\mathrm{k})\ln(1+\exp(\gamma)\mathrm{t}^{\lambda}) \\ \frac{1}{\exp(\gamma)}\left[\exp\left(\frac{-\log(u)-\exp(\alpha' x+k)\ln(1+\exp(\gamma)t_0^{\lambda})+\exp(\alpha' x+k)\exp(\beta_t')\ln(1+\exp(\gamma)t_0^{\lambda})}{\exp(\beta_t')\exp(\alpha' x+k)}\right) - 1\right]^{\frac{1}{\lambda}}, \\ \text{if} \quad -\log(\mathrm{u}) \geq \exp(\alpha'\mathrm{x}+\mathrm{k})\ln(1+\exp(\gamma)\mathrm{t}^{\lambda}) \end{cases} \tag{10}$$

where $u \sim U(0,1)$.

Proposition 5. *Let t be a random variable which follows a Gompertz distribution having scale parameter λ and a shape parameter γ for a dichotomous case of $t < t_0$, with hazard model $m(t \mid x(t), z) = \lambda \exp(\gamma t)\exp(\alpha' x + k)$. Then the cumulative hazard function is $M(t \mid x(t), z) = \frac{\lambda \exp(\alpha' x+k)}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg]$, with inverse cumulative hazard as $M^{-1}(t \mid x(t), z) = \frac{1}{\gamma}\log\left(\frac{\gamma t}{\lambda \exp(\alpha' x+k)} + 1\right)$ if $\mathrm{t} < \frac{\lambda\exp(\alpha'\mathrm{x}+\mathrm{k})}{\gamma}\left(\exp(\gamma\mathrm{t}) - 1\right)$.*

Proof of Proposition 5. We have

$$\begin{aligned} M(t \mid x(t), z) &= \int_0^t \lambda exp(\alpha' x + k)\exp(\gamma u)du \\ &= \lambda exp(\alpha' x + k)\int_0^t \exp(\gamma u)du \\ &= \lambda exp(\alpha' x + k)\left[\frac{1}{\gamma}\exp(\gamma u)\right]_0^t \end{aligned}$$

$$M(t \mid x(t), z) = \frac{\lambda\exp(\alpha' x + k)}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg] \tag{11}$$

Also, the inverse of the cumulative hazard function when $M(t \mid x(t), z) < \frac{\lambda\exp(\alpha' x+k)}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg]$ is:

$$M(t \mid x(t), z) = \frac{\lambda\exp(\alpha' x + k)}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg]$$

$$\frac{\gamma M(t \mid x(t), z)}{\lambda\exp(\alpha' x + k)} = \exp(\gamma t) - 1$$

$$\frac{\gamma M(t \mid x(t), z)}{\lambda\exp(\alpha' x + k)} + 1 = \exp(\gamma t)$$

$$t = \frac{1}{\gamma}\log\left(\frac{\gamma M(t \mid x(t), z)}{\lambda\exp(\alpha' x + k)} + 1\right)$$

$$M^{-1}(t \mid x(t), z) = \frac{1}{\gamma}\log\left(\frac{\gamma t}{\lambda\exp(\alpha' x + k)}+1\right), \ \ if \ \ t < \frac{\lambda\exp(\alpha' x + k)}{\gamma}\Bigg[\exp(\gamma t)-1\Bigg]. \quad (12)$$

Equations (11) and (12) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t < t_0$.

Proposition 6. *Let t be a random variable which follows a Gompertz distribution having scale parameter λ and a shape parameter γ for a dichotomous case of $t \geq t_0$, with hazard model $m(t \mid x(t), z) = \lambda\exp(\gamma t)\exp(\alpha' x + k)$. Then the cumulative hazard function is $M(t \mid x(t), z) = \frac{\lambda\exp(\alpha' x+k)}{\gamma}\Bigg[\exp(\gamma t_0) - 1 + exp(\beta_t')\exp(\gamma t) - exp(\beta_t')\exp(\gamma t_0)\Bigg]$, with inverse cumulative hazard*

$$M^{-1}(t \mid x(t), z) \ = \ \frac{1}{\gamma}\log\left(\frac{\gamma t}{\lambda\exp(\alpha' x + \beta_t' + k)} - \frac{\exp(\gamma t_0) - 1 - \exp(\gamma t_0 + \beta_t)}{\exp(\beta_t)}\right).$$

Proof of Proposition 6. We have

$$\begin{aligned}
M(t \mid x(t), z) \ &= \ \int_0^t \lambda\exp(\alpha' x + \beta_t' y(u) + k)\exp(\gamma u)du \\
&= \int_0^t \lambda\exp(\alpha' x + \beta_t' y(u) + k)\exp(\gamma u)du \\
&= \int_0^t \lambda\exp(\alpha' x + k) exp(\beta_t' y(u))\exp(\gamma u)du \\
&= \lambda\exp(\alpha' x + k)\Bigg[\int_0^{t_0}\exp(\beta_t' y(u))\exp(\gamma u)du + \int_{t_0}^t \exp(\beta_t' y(u))\exp(\gamma u)du\Bigg] \\
&= \lambda\exp(\alpha' x)\Bigg[\int_0^{t_0}\exp(\gamma u)du + \int_{t_0}^t exp(\beta_t' y(u))\exp(\gamma u)du\Bigg]
\end{aligned}$$

Note that $y(u) = 0$ at interval $(0, t_0)$ and $y(u) = 1$ at interval (t_0, t)
$M(t \mid x(t), z) = \lambda\exp(\alpha' x + k)\Bigg[\frac{1}{\gamma}\exp\Bigg[\gamma u\Bigg]_0^{t_0} + exp(\beta_t')\frac{1}{\gamma}\exp\Bigg[\gamma u\Bigg]_{t_0}^{t}\Bigg]$

$$M(t \mid x(t), z) = \frac{\lambda\exp(\alpha' x + k)}{\gamma}\Bigg[\exp(\gamma t_0) - 1 + exp(\beta_t')\exp(\gamma t) - exp(\beta_t')\exp(\gamma t_0)\Bigg] \quad (13)$$

Also, the inverse of the cumulative hazard function when
$M(t \mid x(t), z) \geq \frac{\lambda\exp(\alpha' x+k)}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg]$ is:

$M(t \mid x(t), z) = \frac{\lambda \exp(\alpha' x+k)}{\gamma}\left[\exp(\gamma t_0) - 1 + exp(\beta_t') \exp(\gamma t) - exp(\beta_t') \exp(\gamma t_0)\right]$

$\gamma M(t \mid x(t), z) - \lambda \exp(\alpha' x + k + \gamma t_0) + \lambda \exp(\alpha' x + k) + \lambda \exp(\alpha' x + k + \gamma t_0 + \beta_t') = \lambda \exp(\alpha' x + k + \beta_t' + \gamma t)$

$$\exp(\gamma t) = \frac{\gamma M(t \mid x(t), z) - \lambda \exp(\alpha' x + k + \gamma t_0) + \lambda \exp(\alpha' x + k) + \lambda \exp(\alpha' x + k + \gamma t_0 + \beta_t')}{\lambda \exp(\alpha' x + k + \beta_t')}$$

$$\gamma t = \log\left(\frac{\gamma M(t \mid x(t), z) - \lambda \exp(\alpha' x + k + \gamma t_0) + \lambda \exp(\alpha' x + k) + \lambda \exp(\alpha' x + k + \gamma t_0 + \beta_t')}{\lambda \exp(\alpha' x + k + \beta_t')}\right)$$

$$t = \frac{1}{\gamma}\log\left(\frac{\gamma M(t \mid x(t), z) - \lambda \exp(\alpha' x + k + \gamma t_0) + \lambda \exp(\alpha' x + k) + \lambda \exp(\alpha' x + k + \gamma t_0 + \beta_t')}{\lambda \exp(\alpha' x + k + \beta_t')}\right)$$

$$= \frac{1}{\gamma}\log\left(\frac{\gamma t - \lambda \exp(\alpha' x + \gamma t_0 + k) + \lambda \exp(\alpha' x + k) + \lambda \exp(\alpha' x + \gamma t_0 + \beta_t' + k)}{\lambda \exp(\alpha' x + \beta_t' + k)}\right)$$

$$M^{-1}(t \mid x(t), z) = \frac{1}{\gamma}\log\left(\frac{\gamma t}{\lambda \exp(\alpha' x + \beta_t' + k)} - \frac{\exp(\gamma t_0) - 1 - \exp(\gamma t_0 + \beta_t)}{\exp(\beta_t)}\right) \tag{14}$$

Equations (13) and (14) give the cumulative hazard function and inverse cumulative hazard function respectively for a dichotomous case of $t \geq t_0$.

Hence, the Cox-Gompertz survival time is

$$\mathrm{T} = \begin{cases} \frac{1}{\gamma}\log\left(\frac{\gamma(-\log(\mathrm{u}))}{\lambda\exp(\alpha'\mathrm{x}+\mathrm{k})} + 1\right), & \text{if} \quad -\log(\mathrm{u}) < \frac{\lambda\exp(\alpha'\mathrm{x}+\mathrm{k})}{\gamma}\left[\exp(\gamma\mathrm{t}) - 1\right] \\ \frac{1}{\gamma}\log\left(\frac{\gamma - \log(\mathrm{u})}{\lambda\exp(\alpha'\mathrm{x}+\beta_t'+\mathrm{k})} - \frac{\exp(\gamma\mathrm{t}_0 - 1 - \exp(\gamma\mathrm{t}_0 + \beta_\mathrm{t})}{\exp(\beta_\mathrm{t})}\right), & \text{if} \quad -\log(\mathrm{u}) \geq \frac{\lambda\exp(\alpha'\mathrm{x}+\mathrm{k})}{\gamma}\left[\exp(\gamma\mathrm{t}) - 1\right] \end{cases} \tag{15}$$

where $u \sim U(0, 1)$.

3. The Models

Consider the Cox proportional hazards model Cox (1972) given by

$$m(t \mid x) = m_0(t)h(\mathbf{X}) \tag{16}$$

where $m_0(t)$ is a non-parametric baseline hazard function at time t, $h(\mathbf{X}) = exp(\alpha' x_j)$, and $x_j = (x_{j1}, ...x_{jp})$ is a vector of time invariant co-variates for j^{th} individuals and $\alpha = (\alpha_1, \alpha_2, ...\alpha_p)$ is the unknown parameter vector. The $\exp(\alpha' x_j)$ is a parametric exponential function which assumes parametric forms of the predictors on the hazard models and ensures that the hazard rate remains positive for all values of x. One of the assumptions of the model (16) is that the baseline hazard is common to all the individuals in the population. Extending equation (16) to incorporate a time-varying co-variate $Y_j(t) = Y_1(t), ...Y_q(t)$ which are functions of time or the interaction of co-variates Y_j with a time function $g(t)$. The co-variates under the consideration of non-proportional hazard models is then given by

$$w(t) = f(X_1...X_p, Y_1(t)...Y_q(t))$$

Incorporating time-varying co-variate to equation (16), we have

$$m(t \mid w(x_j, y_j(t))) = m_0(t)\exp(\alpha' x_j + \beta_t' y_j(t)) \tag{17}$$

where α' and β' are coefficients of time-fixed and time-varying co-variates respectively. The model (17) is known as the Extended Cox model or non-proportional hazard model. The choice of $y(t)$ to be used is depends on the choice of the researcher, as stated in Austin (2012). The function of distribution of time-varying co-variate from treated to untreated in medical drug administration is used, for instance,

$$\text{y(t)} = \begin{cases} 1: & if \quad t \geq t_0 \\ 0: & if \quad t < t_0 \end{cases}$$

where t_0 represent the changing point.

Consider a data set of n-subjects from a given population with ith clusters $i = 1, 2, ...k$. Each cluster consists of $n \geq 1$ species which have dependent event times due to unobserved frailty z_j. Introducing frailty term z_j into equation (17) gives

$$m_{ij}(t \mid x(t), z_j, \alpha, \beta_t) = m_0(t) z_j \exp(\alpha' x_{ij} + \beta_t' y(t_{ij})) \tag{18}$$

Equation (18) is a non-proportional hazard model and it can be explained as the conditional hazard function given the observed co-variates and the unobserved co-variates.

$$S_{ij}(t \mid x(t), z_j, \alpha, \beta_t) = \exp(-M_0(t)) z_j \exp(\alpha' x_{ij} + \beta_t' y(t_{ij})) \tag{19}$$

and equation (19) is the survival function given the observed co-variates and the unobserved co-variates.

4. Maximum Likelihood of the Model

The likelihood function for a right censoring is given by:

$$L = \prod_{j=1}^{n} (f_j(t))^{\delta_j} (S_j(t))^{1-\delta_j}$$

where $f_j(t)$, $S_j(t)$ and δ_j is the probability density function, survival function and censoring indicator respectively.
The likelihood function for the j^{th} subject in the i^{th} cluster is illustrated as :

$$L_i = \prod_{j=1}^{n_i} (m_{ij}(t))^{\delta_{ij}} (S_{ij}(t)) \tag{20}$$

where $\delta_{ij} = I(T_i \leq C_i)$ is the censoring indicator which tells us whether an observation is observed or censored.
Hence, using (18) and (19) in (20), the conditional likelihood function for the i^{th} cluster is given by:

$$L_i(\psi, \alpha, \beta \mid z_i) = \prod_{j=1}^{n_i} (m_0(t_{ij}) z_j \exp(\alpha' x_{ij} + \beta_t' y_{ij}(t)))^{\delta_{ij}} \exp(-M_0(t_{ij}) z_j \exp(\alpha' x_{ij} + \beta_t' y_{ij}(t)) \tag{21}$$

where ψ is a vector of parameters of the baseline hazard, α and β are vectors of time-fixed and time-varying co-variates parameters.

$$L_{marg} = \prod_{j=1}^{n_i} \int_0^{\infty} (m_0(t_{ij}) z_j \exp(\alpha' x_{ij} + \beta_t' y_{ij}(t)))^{\delta_{ij}} \exp(-M_0(t_{ij}) z_j \exp(\alpha' x_{ij} + \beta_t' y_{ij}(t)) f_k(z_j) dz \tag{22}$$

where $f_k(z_i)$ is the probability density function of frailties. Integrating out the frailty variable gives the marginal likelihod function.
Suppose the frailty variable z_j follows Gamma distribution then the Probability density function is given as:

$$f_1(z_j) = \frac{z_j^{\frac{1}{\theta}-1} \exp(\frac{-z_j}{\theta})}{\Gamma(\frac{1}{\theta}) \theta^{\frac{1}{\theta}}}, \quad z_j > 0, \quad \theta > 0 \tag{23}$$

Also, suppose the frailty random variable z_j follows an Inverse Gaussian distribution, then Probability density function is given as:

$$f_2(z_j) = \sqrt{\frac{1}{2\pi\theta}}\, z_j^{\frac{-3}{2}} \exp -\frac{(z_j - 1)^2}{2 z_j \theta} \quad z_j > 0, \quad \theta > 0 \tag{24}$$

where θ is the variance of the frailty distribution that measures the heterogeneity which is of interest. Weibull, Log-logistic and Gompertz are used as baseline distributions showing that our model is purely parametric. Recall that our approach is to consider a time-varying co-variate with dichotomous function as stated above. Then from the marginal likelihood (22) above, we obtain the marginal log-likelihood function for $t < t_0$ and $t \geq t_0$. When $t < t_0$, the resulting marginal log-likelihood is the same for when we consider no time-varying function i.e for $t < t_0$, $y(t) = 0$ and for gamma frailty we have

$$\log L_{marg} = \sum_{i=1}^{G}\Big[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}M_0(t_{ij})\exp(\alpha' x)) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \log m_0(t_{ij}))\Big] \quad (25)$$

where $m_0(t_{ij})$ and $M_0(t_{ij})$ is the baseline hazard and cumulative baseline hazard respectively for the baseline distributions used. Also, for $t \geq t_0$, the marginal log-likelihood function for Gamma frailty model is given as:

$$\log L_{marg} = \sum_{i=1}^{G}\Big[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}M_0(t_{ij})\exp(\alpha' x+\beta_t')) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \beta_t' + \log m_0(t_{ij}))\Big] \quad (26)$$

Using **Weibull** as baseline distribution, $m_0(t) = \lambda\gamma t^{v\gamma-1}$ and $M_0(t) = \lambda t^{\gamma}$ can be substituted into equation (25) which is the corresponding marginal log-likelihood function for Gamma frailty model when $t < t_0$ is

$$l(\lambda,\gamma,\theta,\alpha') = \sum_{i=1}^{G}\Big[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}\lambda t^{\gamma}\exp(\alpha' x)) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \log(\lambda\gamma t^{\gamma-1})\Big] \quad (27)$$

Also, for $t \geq t_0$, the marginal log-likelihood for Gamma frailty model is given as:

$$l(\lambda,\gamma,\theta,\alpha',\beta_t) = \sum_{i=1}^{G}\Big[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}\lambda t^{\gamma}\exp(\alpha' x+\beta_t)) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \beta_t + \log(\lambda\gamma t^{\gamma-1})\Big] \quad (28)$$

For the **Log-logistic distribution** having the baseline hazard of $m_0(t) = \frac{\exp(\gamma)\lambda t^{\lambda-1}}{1+\exp(\gamma)t^{\lambda}}$ and $M_0(t) = \ln(1+\exp(\gamma)t^{\lambda})$, when $t < t_0$ can be substituted into (25). The corresponding marginal log-likelihood function for Gamma frailty model when $t < t_0$ is

$$l(\lambda,\gamma,\theta,\alpha') = \sum_{i=1}^{G}\Bigg[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}\ln(1+\exp(\gamma)t^{\lambda})\exp(\alpha' x)) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \log\left(\frac{\exp(\gamma)\lambda t^{\lambda-1}}{1+\exp(\gamma)t^{\lambda}}\right)\Bigg] \quad (29)$$

Also, for t≥ t_0, the marginal log-likelihood function for Gamma frailty model is given as:

$$l(\lambda,\gamma,\theta,\alpha',\beta_t) = \sum_{i=1}^{G}\Bigg[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log(1+\theta\sum_{j=1}^{n}\ln(1+\exp(\gamma)t^{\lambda})\exp(\alpha' x+\beta_t)) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \beta_t + \log\left(\frac{\exp(\gamma)\lambda t^{\lambda-1}}{1+\exp(\gamma)t^{\lambda}}\right)\Bigg] \quad (30)$$

Also, for the **Gompertz distribution** having the baseline hazard of $m_0(t) = \lambda\exp(\gamma t)$ and $M_0(t) = \frac{\lambda}{\gamma}\Bigg[\exp(\gamma t) - 1\Bigg]$, when $t < t_0$, we substitute into (25) which is the corresponding marginal log-likelihood function for Gamma frailty model that is;

$$l(\lambda,\gamma,\theta,\alpha') = \sum_{i=1}^{G}\Bigg[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log\left(1+\theta\sum_{j=1}^{n}\frac{\lambda}{\gamma}\Big\{\exp(\gamma t)-1\Big\}\exp(\alpha' x)\right) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \log\Big(\lambda\exp(\gamma t)\Big)\Bigg] \quad (31)$$

Also, for $t \geq t_0$ the marginal log-likelihood function for Gamma frailty model is given as:

$$l(\lambda,\gamma,\theta,\alpha',\beta_t') = \sum_{i=1}^{G}\Bigg[d_i\log\theta - \log\Gamma(\frac{1}{\theta}) - (\frac{1}{\theta}+d_i)\log\left(1+\theta\sum_{j=1}^{n}exp(\alpha' x+\beta_t)\frac{\lambda}{\gamma}\Big\{\exp(\gamma t)-1\Big\}\right) + \log\Gamma(\frac{1}{\theta}+d_i) + \sum_{j=1}^{n_i}\delta_{ij}(\alpha' x + \beta_t + \log(\lambda\exp(\gamma t))\Bigg] \quad (32)$$

The maximum likelihood estimates of the proposed models in equation (27),(28), (29), (30), (31) and (32) can be obtained by using optimization method of Newton-Raphson iterative procedure for parameters of interest λ, γ, θ, α' and β_t' setting the first derivatives of the interested parameters to zero and solve the equation simultaneously.

INVERSE GAUSSIAN FRAILTY MODELS

At $t < t_0$, $y(t) = 0$, for Inverse Gaussian frailty we have the marginal log-likelihood as

$$\begin{aligned} l_{marg} = \sum_{i=1}^{G} \Bigg[d_i \sum_{j=1}^{n} \Big(\alpha' x + \log(m_0(t_{ij}) \Big) - \frac{1}{2} \log \pi - \frac{1}{2} \log 2\theta \\ - \frac{2\Big(1 + 2\theta \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\alpha' x}\Big)^{\frac{1}{2}}}{\theta} + \left(\frac{2}{2d_i - 3}\right) \log \pi + \left(\frac{2}{2d_i - 3}\right) \log \theta \\ - \left(\frac{2}{2d_i - 3}\right) \log \left(2\Big(1 + 2\theta \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\alpha' x}\Big)^{\frac{1}{2}} \right) \Bigg] \end{aligned} \quad (33)$$

At $t \geq t_0$, $y(t) = 1$, for Inverse Gaussian frailty we have the marginal likelihood as

$$\begin{aligned} l_{marg} = \sum_{i=1}^{G} \Bigg[d_i \sum_{j=1}^{n} \Big(\alpha' x + \beta_t + log(m_0(t_{ij}) \Big) - \frac{1}{2} \log \pi - \frac{1}{2} \log 2\theta \\ - \frac{2\Big(1 + 2\theta \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\alpha' x + \beta_t}\Big)^{\frac{1}{2}}}{\theta} \\ + \left(\frac{2}{2d_i - 3}\right) \log \pi + \left(\frac{2}{2d_i - 3}\right) \log \theta \\ - \left(\frac{2}{2d_i - 3}\right) \log \left(2\Big(1 + 2\theta \sum_{j=1}^{n_i} M_0(t_{ij}) e^{\alpha' x + \beta_t}\Big)^{\frac{1}{2}} \right] \end{aligned} \quad (34)$$

The baseline hazards and cumulative baseline hazard functions for each of the aforementioned baseline distributions can be substituted into equation (33) and (34) and the corresponding marginal log-likelihood are obtained and resolved for parameter estimates using Newton-Raphson iterative procedure.

5. Application to data on Lung Cancer

The methods discussed above are applied to a study conducted by the US Veterans Administration on male patients with advanced inoperable lung cancer under a standard therapy or a test chemotherapy. Time to death was recorded for 137 patients, while 9 withdrew from the study before death. Various co-variates were also documented for each patient these include:

(i) The type of lung cancer treatment, denoted as Trt;

$$\text{Trt} = \begin{cases} 1, & \text{if} \quad \text{standard;} \\ 2, & \text{if} \quad \text{test} \quad \text{drug.} \end{cases}$$

(ii) The type of cell involved, denoted as cell-type;

$$\text{cell} - \text{type} = \begin{cases} 1, & \text{if} \quad \text{squamous;} \\ 2, & \text{if} \quad \text{smallcell;} \\ 3, & \text{if} \quad \text{adeno;} \quad \text{and} \\ 4, & \text{if} \quad \text{large.} \end{cases}$$

(iii) The survival time (in days) since commencement of the treatment is denoted as survival time;
(iv) The status of the patient, denoted as status;

$$\text{status} = \begin{cases} 1, & \text{if} \quad \text{dead;} \\ 2, & \text{if} \quad \text{alive.} \end{cases}$$

(v) Karnofsky performance or score denoted as karno;
(vi) The time (in months) since diagnosis, denoted as diagtime;
(vii) The age (in years) of the patients, denoted as age; and
(viii) Prior therapy, denoted as prior;

$$\text{prior} = \begin{cases} 0, & \text{if} \quad \text{none;} \\ 10, & \text{if} \quad \text{yes.} \end{cases}$$

This data set was extracted from Kalbfleisch and Prentice (2002) in which a Weibull regression model was fitted with 8 regressor variables In the weibull regression model fitted, the model captures only the observed co-variates and treats all the co-variates as time-fixed meanwhile this present study captures both the observed and unobserved co-variates with the help of the frailty parameter introduced to the Cox model and these co-variates were treated based on time-fixed and time-varying concurrently.
The frequentist approach is used for the estimation of parameters in this paper. The parametric frailty model (parfm) package was used as a means of estimation of parameters in this study. It is a recent R package discussed by Rotolo *et al.* (2012) that works for the gamma, the inverse Gaussian, and the positive stable proportional hazards frailty models with baseline distributions like exponential, Weibull, Gompertz, Log-normal, or Log-logistic distributions. The main merit of this package *parfm* depends on the selection of frailty distributions and parametric baseline hazards it works for. The estimation of the parameters is carried out by maximizing the marginal log-likelihood.

Table 1. Cox regression model

co-variates	coef	exp(coef)	se(coef)	Z	P-value
trt	0.19305	1.213	0.18645	1.035	0.300
diagtime	0.00172	1.002	0.00900	0.191	0.85
age	-0.00388	0.996	0.00925	-0.420	0.670
karno	-0.03408	0.966	0.00534	-6.381	1.8e-10
prior	-0.00776	0.992	0.02215	-0.350	0.73

Table 2. Checking for PH assumption

co-variates	rho	chisq	p
trt	-0.0929	1.18	0.278
diagtime	0.1411	2.50	0.114
age	0.1855	5.12	0.0237
karno	0.3582	16.49	0.0000489
prior	-0.1832	4.58	0.0323
Global	NA	21.08	0.000784

Table 3. Extended Cox Model

co-variates	rho	chisq	p
trt	-0.00403	0.00222	0.962
diagtime	0.02446	0.05322	0.818
age(y(t))	0.13117	1.75537	0.185
karno(y(t))	0.09022	0.97421	0.324
prior(y(t))	-0.00959	0.01263	0.911
Global	NA	5.74653	0.332

The change point was located at the median survival time which is 80 years. Preferably, it is better to use the median over the mean due to its popularity in describing the central tendency in survival time data because of its skewness. The time varying co-variate $Y_i(t)$ was thus generated as

$$\text{y(t)} = \begin{cases} 1 & \text{for} \quad \text{t} \geq 80 \\ 0 & for \quad t < 80 \end{cases}$$

6. Discussion

Table 1 shows the output of the Cox regression model with the following variables; treatment (trt), age of the patients, diagtime, karnofsky performance and prior

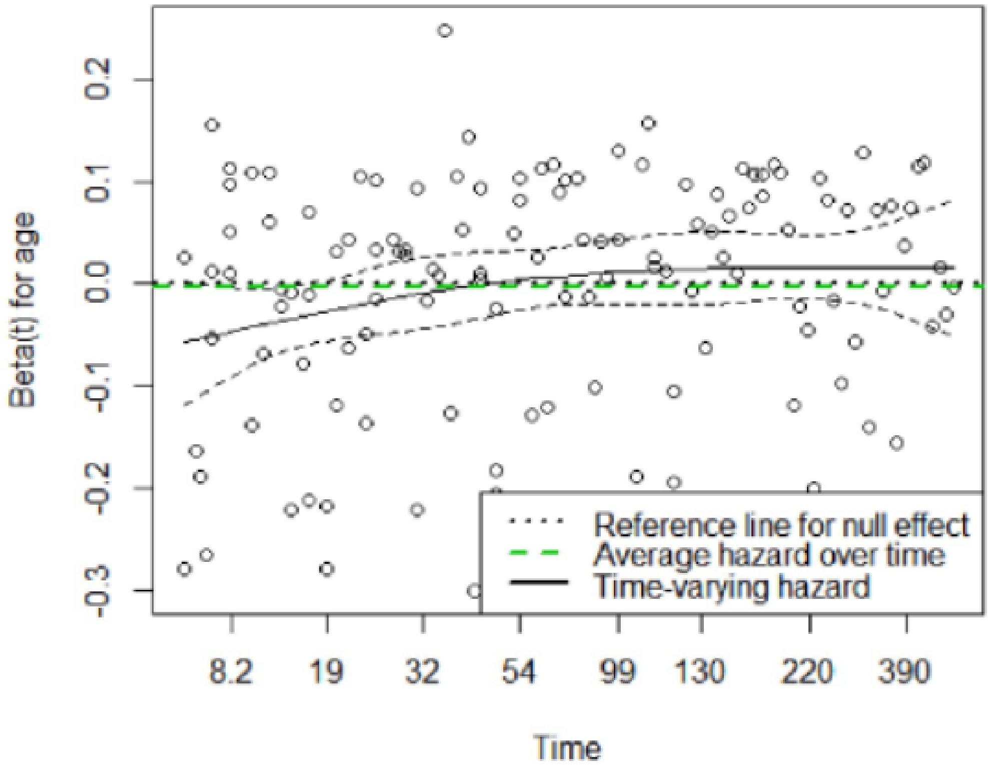

Fig. 1. Schoenfeld Residual Plot for Age of patients

Table 4. Parameters Estimation Using Weibull as baseline when all co-variates were treated as time-fixed

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	0.148(0.130)	0.170(0.161)
γ	1.055(0.071)	1.055(0.071)
λ	0.051(0.040)	0.051(0.040)
trt (SE) (p-value)	0.240(0.198)(0.224)	0.233(0.198)(0.240)
diagtime (SE) (p-value)	0.001(0.009)(0.930)	0.001(0.009)(0.919)
age (SE) (p-value)	-0.005(0.009)(0.553)	-0.006(0.009)(0.552)
karno (SE) (p-value)	-0.033(0.005)(0.000)	-0.033(0.005)(0.000)
prior(SE) (p-value)	0.001(0.023)(0.978)	0.001(0.023)(0.976)

therapy. It gives the result of the coefficient, hazard rates, the standard error of the coefficient, Wald(Z) and the p-value. From the output of the Cox regression model in Table 1 the result shows that the p-value for karnofsky performance score is highly significant (1.8e-10), thus is a sign that the variable is time-varying.

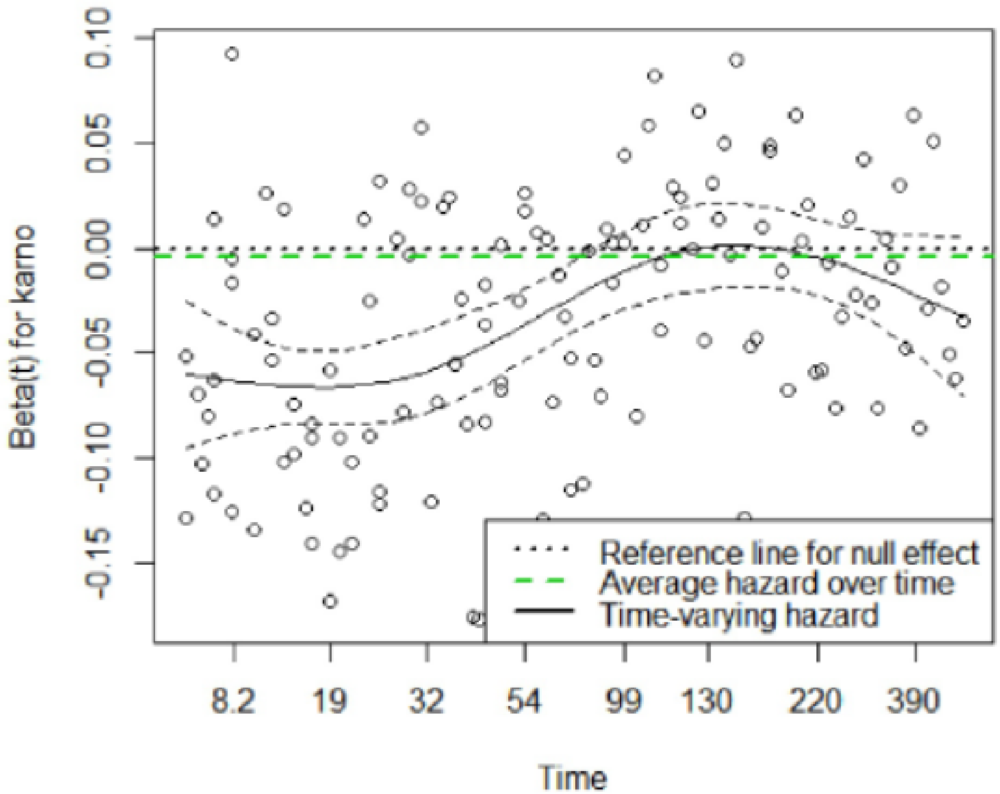

Fig. 2. Schoenfeld Residual Plot for Karnofsky Performance

Table 5. Parameters Estimation Using Weibull as baseline when time-varying co-variates were considered

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	0.022(0.046)	0.022(0.048)
γ	1.468(0.100)	1.468(0.100)
λ	0.006(0.003)	0.006(0.003)
trt (SE) (p-value)	-0.033(0.190)(0.861)	-0.036(0.190)(0.852)
diagtime (SE) (p-value)	0.006(0.007)(0.336)	0.006(0.007)(0.334)
age(y(t)) (SE) (p-value)	-0.020(0.008)(0.012)	-0.020(0.008)(0.012)
karno(y(t)) (SE) (p-value)	-0.024(0.007)(0.001)	-0.024(0.007)(0.001)
prior(y(t)) (SE) (p-value)	-0.048(0.030)(0.109)	-0.048(0.030)(0.108)

From the output in Table 2 the result shows that there are significant deviations from the proportional hazards assumption for the following variables: age of patients (p=0.0237), Karnofsky performance $(p = 0.0000489)$ and prior therapy $(p = 0.0323)$. These p-values showed that the age of patients, Karnofsky per-

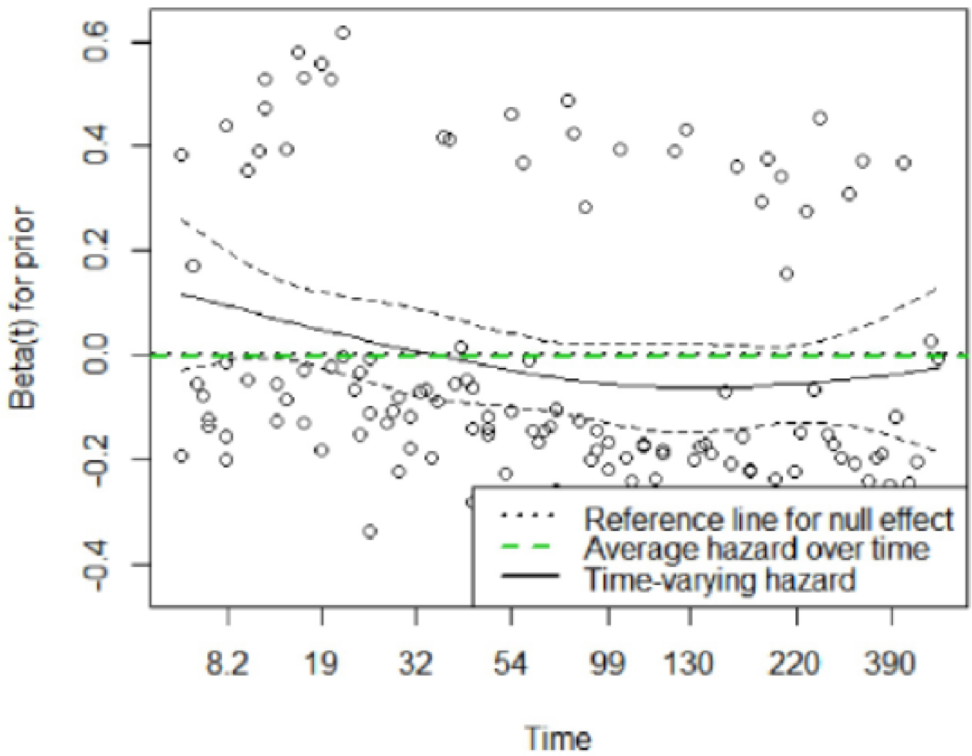

Fig. 3. Schoenfeld Residual Plot for Prior Theraphy

Table 6. Parameter Estimation Using Log-logistic as baseline when all coivariates are treated as time-fixed

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	0.119(0.136)	0.114(0.121)
γ	-6.341(0.468)	-6.341(0.468)
λ	1.276(0.120)	1.277(0.120)
trt (SE) (p-value)	0.418(0.194)(0.032)	0.416(0.195)(0.033)
diagtime (SE) (p-value)	0.004(0.009)(0.655)	0.004(0.009)(0.653)
age (SE) (p-value)	0.020(0.007)(0.005)	0.021(0.007)(0.005)
karno (SE) (p-value)	-0.021(0.005)(0.000)	-0.021(0.005)(0.000)
prior (SE) (p-value)	0.008(0.023)(0.739)	0.008(0.023)(0.736)

formance and prior therapy are time-varying. In general, an associated global significant test gives a p-value of (0.000784) which is an indication of lack of fit of the model. Having discovered from Table 2 that there is an indication of time varying co-variates in variables used, the Table 3 then shows the result from an extension of the Cox model known as Extended Cox model or Non-proportional

Table 7. Parameter Estimation Using Log-logistic as baseline when time-varying co-variates were considered

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	12.251(10.307)	11.238(11.209)
γ	-9.214(0.652)	-9.199(0.655)
λ	1.686(0.136)	1.688(0.136)
trt (SE) (p-value)	0.319(0.209)(0.128)	0.332(0.212)(0.118)
diagtime (SE) (p-value)	0.005(0.007)(0.412)	0.005(0.007)(0.415)
age(y(t)) (SE) (p-value)	-0.021(0.008)(0.010)	-0.021(0.008)(0.011)
karno(y(t)) (SE) (p-value)	-0.019(0.007)(0.006)	-0.019(0.007)(0.007)
prior(y(t)) (SE) (p-value)	-0.029(0.029)(0.311)	-0.029(0.029)(0.314)

Table 8. Parameters Estimation Using Gompertz as baseline when all co-variates were treated as time-fixed

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	1.160(1.474)	3.957(1.875)
γ	0.000(0.000)	0.001(0.000)
λ	0.043(0.039)	0.045(0.038)
trt (SE) (p-value)	-0.112(0.200)(0.574)	-0.118(0.201)(0.570)
diagtime (SE) (p-value)	-0.001(0.011)(0.952)	-0.001(0.011)(0.944)
age(y(t))(SE) (p-value)	0.003(0.009)(0.771)	0.002(0.009)(0.758)
karno(y(t)) (SE) (p-value)	-0.022(0.005)(0.000)	-0.020(0.005)(0.000)
prior(y(t)) (SE) (p-value)	-0.014(0.024)(0.542)	-0.013(0.024)(0.540)

Table 9. Parameter Estimation Using Gompertz as baseline when time-varying co-variates were considered

Parameter	Gamma Frailty	Inverse Gaussian Frailty
θ(SE)	2.250	5.011
γ	0.003(0.001)	0.005(0.000)
λ	0.278	1.396
trt (SE) (p-value)	-0.369(0.207)(0.075)	-0.654(0.213)(0.002)
diagtime (SE) (p-value)	0.014(0.006)(0.028)	0.020(0.006)(0.001)
age(y(t))(SE) (p-value)	-0.022(0.008)(0.008)	-0.004(0.007)(0.609)
karno(y(t)) (SE) (p-value)	-0.015(0.007)(0.034)	-0.040(0.007)(0.000)
prior(y(t))(SE) (p-value)	-0.084(0.035)(0.015)	-0.038(0.029)(0.192)

hazards model. Here, in Table 3 it is seen that the model is stable.

Figure 1 shows the effect of time-varying coefficient of age on mortality outcome which varies over time. The horizontal time axis is in transformed scale, which is the default setting in the cox.zph() function. The dashed lines are lower and upper limits of confidence interval of the effect of age. It is noted that the effect of age is time-varying (i.e. time-varying coefficient). The graph in Figure 2 shows the effect of time-varying coefficient of Karnofsky performance on mortality outcome which

Table 10. Akaike Information criterion (AIC) to test goodness-of-fit when co-variates are all treated as time-fixed

	Gamma Frailty	Inverse Gaussian Frailty
Weibull	1458.417	1458.404
Log-logistic	1477.319	1477.318
Gompertz	1470.822	1470.810

Table 11. Akaike Information criterion (AIC) to test goodness-of-fit when time-varying were considered

	Gamma Frailty	Inverse Gaussian Frailty
Weibull	1368.570	1368.577
Log-logistic	1402.577	1405.482
Gompertz	1409.507	1435.293

varies over time based on the evidence shown.It is noted that the effect of karno is time-varying (i.e. time-varying coefficient). Likewise for prior therapy graph in Figure 3, the coefficient is time-varying. Hence, it is concluded that the three co-variates and coefficients at 95% confidence are varying with time.

From Table 4 to Table 9, it can be seen that the value of the parameter θ is greater than zero this indicates of the presence of an unobserved co-variate which the Cox model cannot fully account for. Also, Table 10 and Table 11 gives the value of the AIC as this compares the fit of the Gamma and Inverse Gaussian frailty model for both the time-fixed coco-variate-variates and the time-varying s respectively. In the time-fixed, the Inverse Gaussian frailty model fit better into the data set due to its smaller value of AIC which is in contrast to the time-varying case as the AIC value of the Gamma frailty model is smaller which makes it more fit to the data set than the Inverse Gaussian Frailty model.

7. Conclusion

The main purpose of this study is to compare the Gamma and Inverse Gaussian frailty models for when co-variates are treated as time-fixed and as time-varying. Real-life data set was used for comparing when all the co-variates are treated as time-fixed and time-varying for the three baseline distributions used. The Inverse Gaussian frailty model fits better than the Gamma frailty model when it is time-fixed due to its smaller value of AIC. Gamma frailty model on the other hand fits better than the Inverse Gaussian frailty model when it is time-varying due to its smaller value of AIC. The study concluded that failure to account for frailty in the Cox model with time-varying co-variates makes the model less flexible such that it may not adequately account for co-variates effect.

Acknowledgment The authors acknowledge and expresses his warm thanks to the editorial team of *Afrika statistika* most especially Professor Gane Samb Lo for scientific supervising and his moral support on this paper. This paper could not have been completed without his careful watch.

References

Abrahamowicz, M., & MacKenzie, T. A. (2007). Joint estimation of time dependent and non-linear effects of continuous co-variates on survival. Statistics in Medicine, 26(2), 392-408. doi:10.1002/sim.2519

Adeleke, K. A., Abiodun, A. A. and Ipinyomi, R. A. (2015). Semi-Parametric Non-Proportional Hazard Model with Time varying co-variate. *Journal of Modern Applied Statistical methods*, 14(2), 68 − 87

Ata, N. & Sozer, M. T. (2007). Cox regression models with nonproportional hazards applied to lung cancer survival data. Hacettepe Journal of Mathematics and Statistics, 36(2), 157-167.

Austin, P. C. (2012). Generating survival times to simulate Cox proportional hazards models with time-varying co-variates. *Statistics in Medicine*, Vol 31 (29), pp. 3946-3958.

Bender. R., Augustin, T., & Blettner, M. (2005). Generating survival times to simulate Cox proportional hazards models. Statistics in Medicine, 24(11), 1713-1723. doi:10.1002/sim.2059

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society B*, 34:187-220

Cox, D. R. (1975). Partial likelihood. Biometrika, 62(2), 269-276 doi:10.2307/2335362.

Duchateau, L. and Janssen, P. (2008). The Frailty models. Springer-Verlag, New York.

Fagbamigbe A.F, Afolabi R.F., Alade K. Y, Adebowale A.S., and Yusuf B.O (2019); Unobserved Heterogeneity in Determinants of Under-five Mortality in Nigeria: Frailty Modeling in Survival Analysis. *African Journal of Applied Statistics*, Vol. 19 (1), pp. 565 - 587.

Henderson, R. and Oman, J. P. (1999). Effect of frailty on marginal regression estimates in survival. *Journal of the Royal Statistical Society*,61:367-380

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. Biometrika, 73:387-396

Hougaard, P.(2000). Analysis of multivariate survival data . Springer, New York

Keyfitz, N. and Littman, G. (1979). Mortality in a heterogeneous population. Population Studies, 33: 333-342

Kronborg, D. and Aaby, P. (1990), "Piecewise Comparison of Survival function in Stratified Proportional Hazards Models", Biometrics, 46, 375-380.

Lancaster,T. (1979). Econometric methods for the duration of unemployment. Econometrica 47:939-956

Lehr, S. (2004). Over-fit in the analysis of time-dependent effects of prognostic factors. Ph.D. Thesis, University of Vienna.

Kalbfleisch, J. D. and Prentice, R. L. (2002). The statistical analysis of Failure Time Data. Wiley, New York.

Lee, E. T. and Go, O. T. (1997). Survival analysis in public health research. Annual Review of Public Health, 18: 105-134

Leemis, L. M. (1987). Technical note-Variate generation for accelerated life and proportional hazards models. Operations Research, 35(6), 892-894. Doi:10.1287/opre.35.6.892

Leemis, L. M., Shih, L. H., & Reynertson, K. (1990). Variate generation for accelerated life and proportional hazards models with time dependent co-variates. Statistics and Probability Letters; 10(6), 335-339. doi:10.1016/0167-7152(90)90052-9

Robert, C. P. and Casella, G. (2009). Introducing Monte Carlo Methods with R, Springer-Verlag

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84:487-493.

Rotolo, F., Munda M., and Legrand, C. (2012). parfm: Parametric Frailty Models. R package version 2.5.2, URL http://CRAN.R-project.org/package=parfm

Shih, L. H. & Leemis, L. M. (1993). Variate generation for a non homogeneous poisson process with time dependent co-variates. Journal of Statistical Computation and Simulation; 44(3-4), 165-186. doi:10.1080/00949659308811457

Sylvester, M. P. & Abrahamowicz, M. (2007). Comparison of algorithms to generate event times conditional on time-dependent co-variates. Statistics in Medicine; 27(14), 2618-2634.doi:10.1002/sim.3092

Zhang, Z., Reinikainen J., Adeleke, K. A., Pieterse, M. E. and Groothuis-Oudshoorn, C. G.(2018) Time-varying co-variates and coefficients in Cox regression models. Ann Transl. Med 2018;6(7):121.

Zhou, M (2001). Understanding the Cox regression models with time change co-variate. The American Statistician; 55(2):153155:10.1198/000313001750358491