Malaysian Journal of Applied Sciences 2019, **Vol 4(1)**: 1-14

eISSN 0127-9246 (Online)

Malaysian Journal of Applied Sciences

ORIGINAL ARTICLE

An Empirical Comparison between Partial Likelihood and Penalized Partial Likelihood Estimators for Semi-parametric Non-proportional Hazards Models with Frailty

Alfred A. Abiodun[a] , Benjamin A. Oyejola[a] and Kazeem A. Adeleke[b]

[a]Department of Statistics, University of Ilorin, PMB 1515, Ilorin, Kwara State, Nigeria.
[b]Deparrment of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria.

*Corresponding author: alfredabiodun1@gmail.com

Received: 04/02/2019, Accepted: 09/04/2019

Abstract

This study compared partial likelihood (PL) and penalized partial likelihood (PPL) estimators in non-proportional hazards model with dichotomous time-varying covariates and subject–specific frailty. We considered Gamma and Inverse Gaussian as frailty distributions. The methods were illustrated with a dataset on diabetes. Extensive numerical studies were conducted using Monte Carlo simulations to compare the efficacy of the methods in terms of Relative Bias (RB) and Root Mean Square Error (RMSE). A sensitivity analysis was carried out to assess the power of the estimators under misspecification of frailty distributions. It was found, that PPL estimator generally outperformed PL estimator in all scenarios considered. Efficiency was found to increase with increase in sample size, and decrease with increase in censoring proportion. The sensitivity analysis conducted to assess the effect of frailty misspecification revealed that sample size, proportion of censored observations and the shape of the frailty distribution (log-skewed) severely affected the power of the estimators.

Keywords: Survival time; non-proportional hazards; frailty; time-dependent covariate; relative bias

Introduction

Non-proportional hazard models have become popular in the analysis of time-to-event data. In practical situations, the proportionality assumption which is often made regarding proportional hazards (PH) models is violated. This implies that the covariate under investigation no longer has a constant impact on the hazard ratio over time, and therefore, applicability of proportional hazards models, proposed by Cox (1972), becomes inappropriate. Correctly accounting for time-varying covariates is important because it allows one to avoid the problem of survivor-treatment or immortal-time bias (Suissa, 2007; Beyersmann, et al., 2008; Austin et al., 2006). Non-proportional hazards can arise if some covariates only affect survival up until sometime t or if the size of their effects change over time. In such a situation, Extended Cox regression model will be more appropriate to model such time-varying covariates instead of the usual standard Cox model. Some of the studies that have been carried out in the recent times in this direction include Zhou (2001), Robert and Casella (2009), Austin (2012 and Adeleke et al. (2015). Valenta and Weissfeld (2002) explicitly simulated a piecewise-exponential model to evaluate Gray's piecewise constant time-varying coefficients

model in a study on problems with vaguely defined disease states. Lee, Seo, and Shin (2011) developed a general method that allow a set of covariates to have heterogeneous effects across the domain of the threshold covariate (Zhang et al., 2014)

Incorporating subject-specific frailty or unobserved heterogeneity, in survival analysis has become popular, especially in medical research, because subjects differ substantially in characteristics and this causes variations in the effects of observed covariates amongst them. This may be as a result of exposure of individuals to different risk levels, which makes them experience failure events than others, even after controlling for the known risk factors. Also, some relevant covariates may be unavailable to the researchers (Munda, 2012), and this makes the assumption of independent and identical distribution, which is often made regarding survival data of different patients implausible.

Frailty model was introduced in the biostatistical literature by Vaupel, Manton, and Stallard (1979), and discussed in details by Hougaard (2000), Duchateau and Janssen (2008), Wienke (2010), and Munda et al. (2012).

The general family of distributions commonly used for frailties is the family of power variance function (PVF) distributions (Hougaard, 2000), and these include gamma, positive stable and inverse Gaussian distributions. They are most popular because their densities are easy to use, as they lead to closed-form expressions for the expectation of the frailties. Recent research have also used an extension of lognormal distribution to log-skewed normal distribution (Azzalini, 1985) as frailty distributions.

The problem of frailty misspecification is actually a part of a more general area of misspecification in the statistical literature. Neuhaus et al. (1992) concluded in his study that the misspecified mixed effects model may result in regression parameter estimates that are asymptotically biased (inconsistent). Pickles and C'rouchley (1995) performed simtlations to examine the effect of frailty misspecification on regression coefficient estimation where the results indicate robustness to misspecification, and this is in contrast to the conclusions of Heckrnan and Singer (1984), who found high sensitivity of parameter estimates from economic duration data to the assumed functional form for the distribution of unobserved variables.

Estimation of parameters in Cox model is often done by maximizing the partial likelihood (PL) or the log partial likelihood either directly (Andersen et al., 1997) or by using the expectation maximization (EM) algorithm (Nielsen et al., 1992; Vaida Xu, 2000; Duchateau & Janssen, 2008). However, the resulting estimates from partial could suffer from substantial bias caused by the presence of many nuisance parameters or may have large variance when collinearity exists among the explanatory variables (Huang and Harrington, 2002). Penalized partial likelihood (PPL) method of estimation (Therneau, Grambsch, and Pankratz, 2003, Therneau and Grambsch, 2000) has in recent times become a popular alternative to Cox PL method for (censored) survival data. Penalized estimating equations have been used in settings such as nonparametric regression modelling, where a penalty term is used to reduce over fitting with high dimensional models, and the key step is to apply a penalty function to smoothen covariate effects of interest. Verweij and Van Houwelingen (1993) have used a cross-validated partial likelihood (CVL) to choose the penalty parameter, where the optimal penalty parameter is often chosen to maximize the sum of the contributions of each subject to the log partial likelihood and thus minimize the CVL function.

This study therefore, aims to comparing the two estimators (PL and PPL) using data on Diabetes and simulated data under the violation of proportional hazards assumptions in the presence of subject-specific frailty, and to assess their sensitivity to frailty misspecification.

Materials and Methods

Model Formulation

Consider the standard Cox proportional hazards model given by

$$\psi_i(t, z_i) = \psi_0(t)\exp(\gamma' Z_i) \quad (1)$$

where is a non-parametric baseline hazard function, Zi is a vector of time-invariant covariates for ith individual and is a p vector of regression coefficients. The model in (1) assumes that the hazard at time t is constant over time. An important feature of the model which concerns the proportional hazards assumption is that the baseline hazard is a function of t, but does not involve the Z's. Often time, the effect of a covariate on survival probability may be time-varying that depends on t, denoted by Xi(t) = Xi*r(t), which is an interaction of covariate Xi with the function of time r(t). The covariates for consideration under non-proportional hazards model framework may then be extended as

$$Zi(t) = (Zi , Xi(t)) \quad (2)$$

We can write the model in (1), including time-varying covariate as

, $$\psi_i(t|Z_i(t)) = \psi_0(t)\exp(\gamma' Z_i + \beta_t' X_i(t)) \quad (3)$$

The function r(t) in the time-varying covariates Xi(t) is usually selected according to the information level of the researcher. In this study, a dichotomous one-step change point function has been considered as in Austin (2012), and this is given as.

, $$r(t) = \begin{cases} 1 \; if \;\; t \geq t_a \\ 0 \; if \;\; t < t_a \end{cases} \quad (4)$$

where ta is the threshold value representing the point where the change occurs. An example of time varying covariates of this form is drug administration to patient on hospital admission which may have different effects at early and later time of the admission. Other common choices of r(t) include t and $\log(t)$.

Incorporation of subject-specific frailty

Let Ti be the event time of the ith subject which is censored by a variable Ci. Then the observed survival time can be given as the minimum of the event time Ti and the censoring time Ci, with the censoring indicator given by δj =I(Ti ≤ Ci) where I(.) is the censoring indication function. If Wj is the associated frailty for subject i, then given Zi(t) = (Zi,Xi(t)) and Wj, the model with subject-specific frailty, extending (3) can be given (Wienke 2010, Munda et al. (2012) as.

$$\psi_i(t|Z_i(t), W_i, \gamma, \beta_t) = W_i\psi_0(t)\exp(\gamma' Z_i + \beta_t' X_i(t)) \quad (5)$$

$$H(t|Z_i(t), W_i) = \int_0^t \psi_0(u)\exp(-\int_0^t \psi(u|Z_i(t), W_i)du) \quad (6)$$

and the conditional survival function is given by

$$S(t|Z_i(t), W_i) = \exp(-H(t|Z_i(t), W_i) \quad (7)$$

The model in (7) represents the probability of being alive at time t given the frailty Wi.

Frailty Distributions

We give three frailty distributions used in this study

Gamma Frailty (GAF) Distribution

The one parameter gamma distribution, has been the most popularly used frailty distribution in survival analysis. The density function is given as

$$h_\theta(w) = \frac{w^{\frac{1}{\theta}-1}\exp(-w/\theta)}{\Gamma(1/\theta)\theta^{\frac{1}{\theta}}}, w,\ \theta > 0 \qquad (8)$$

where Γ (.) is the Gamma function.

For the purpose of identifiability, it is assumed that the random variable W has mean and variance given respectively as $E(W) = 1 \text{ and } Var(W) = \theta.$
Larger values of indicate that there is a higher degree of heterogeneity among subjects.

Inverse Gaussian Frailty (IGF) Distribution

The density for Inverse-Gaussian distribution is given as

$$h_\theta(w) = (2\pi\theta w^3)^{-12}\exp\left\{\frac{1}{20}\left(w - 2 + \frac{1}{w}\right)\right\} \qquad (9)$$

Log-skewed Normal (LSN) Distribution

The log-skewed normal density distribution for a continuous random variable W (Azzalini, 1985), is given as

$$\phi(w,\alpha) = 2\phi(w)\Phi(\alpha,w), \quad -\infty < w < \infty \quad , \quad -\infty < \alpha < \infty \qquad (10)$$

where ϕ and Φ denote the standard normal density and distribution function respectively, and α is the skewness parameter such that α > 0 produces a distribution with positive skewness and α < 0 corresponds to negative skewness; if α = 0, the distribution reduces to the usual standard normal density.
It should be noted that W > 1 indicates that a subject is frail, whereas W < 1 indicates that the subject is strong and has lower risk.

Methods of Estimation
Estimation of parameters in this study is based on the partial likelihood (PL) and penalized partial likelihood (PPL) methods (McGilChrist & Aisbett, 1991, McGilchrist, 1993).

Partial Likelihood (PL)
The likelihood is formed as a combination of the failures and the censored observations as in the standard survival models and is given as.

$$, \qquad L = \prod_{i=1}^{n}[f_i(t|Z_i(t)]^{\delta_i}[S_i(t|Z_i(t)]^{1-\delta_i} \qquad (11)$$

where $S_i(t) = 1 - F_i(t)$ is the survival function of the event time.
Writing the hazard function in terms of the density and survival functions, we have

$$\psi_i(t) = \frac{f_i(t)}{S_i(t)}$$

Then the conditional likelihood for the ith individual given frailty Wi can be written as

$$L_i(\gamma, \beta_t|W_i) = \prod_{i=1}^{n}[W_i\psi_0(t_i)\exp(\gamma'Z_i + \beta_t'X_i(t)]^{\delta_i}\exp[(-H_0(t)W_i\exp(\gamma'Z_i + \beta_t'X_i(t))] \quad (12)$$

where is the cumulative hazard function. Unconditionally, we can obtain the marginal likelihood of the ith individual by integrating out so that we have

$$L_{mag}(\gamma, \beta_t | W_i) = \int_0^{\infty} \prod_{i=1}^{n} [W\psi_0(t_i)\exp(\gamma' Z_i + \beta_t' X_i(t)]^{\delta_i} \times \exp[(-H_0(t)W\exp(\gamma' Z_i + \beta_t' X_i(t))]h_\theta(w)dw \quad (13)$$

Where $h_\theta(w)$ is the probability density function of the frailty given in (8) and (9). The marginal likelihood in (13) is thus maximized to obtain the partial likelihood (PL) estimators $(\hat{\gamma}, \hat{\beta}_t, \hat{\theta})$.

Penalized Partial Likelihood (PPL)

Rather than integrating out the frailty i as in (13), the penalized partial likelihood method treats the frailty density in the complete data likelihood as a penalty term. Therefore, following (McGilchrist, 1993), the penalized partial likelihood for the frailty model can be written as

$$PPL(\gamma, \beta_t | W_i) = \log\left[\left(\frac{W_i \exp\{\gamma' Z_i + \beta_t' X_i(t)}{\sum_{R_t} W_s \exp\{\gamma' Z_s + \beta_t' X_s(t)} + \sum_{i=1}^{n} log(h_\theta(w)\right)^{\delta_i}\right] \quad (14)$$

,

which is maximized over both (γ, β_t) and Wi to obtain the corresponding penalized partial likelihood estimators.

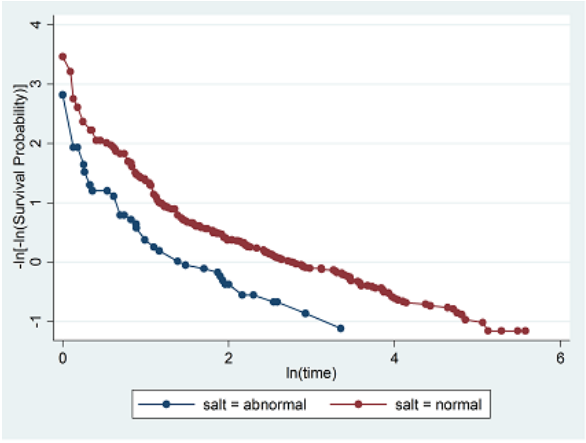
Maximization of the penalized partial likelihood criterion leads to the same parameter estimates and the frailties Wi as in the EM algorithm (Therneau et al., 2003). Using the frailty distributions in (8) and (9) with the likelihoods in (13) and (14), four model configurations were considered in this study which are: Partial likelihood with Gamma frailty (PL-GAF), Partial likelihood with Inverse Gaussian frailty (PL-IGF), Penalized Partial likelihood with Gamma frailty (PPL-GAF) and Penalized Partial likelihood with Inverse Gaussian frailty (PPL-IGF).

Results

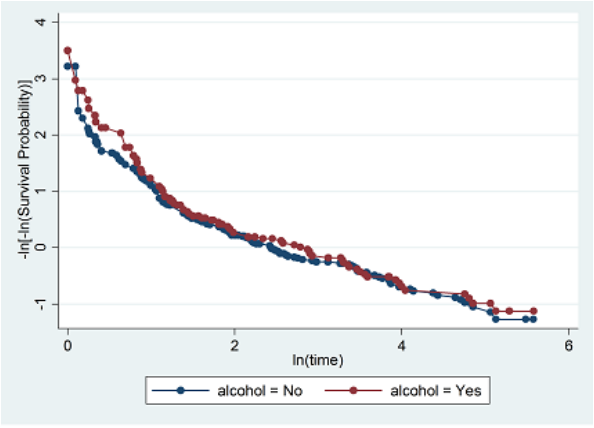
Application to Data on Diabetic Patients

Data were collected on 150 diabetic patients who were diagnosed and admitted at the University of Maiduguri Teaching Hospital, Borno State. From these, 60 patients died during the duration of admission and the remaining 90 were either lost to follow-up or were still alive at the time of data collection and were censored at the time they were last seen or time of data collection. Survival time was defined as the time to death due to diabetes, recorded in weeks. The covariates thought to be associated with diabetes were also collected, and these included: age of the patient in years, sex (male=1, female=0), body mass index in kg/m2, alcohol intake (yes=1, no=0), salt intake (normal=1, abnormal=0), family history (inherited=1, uninherited=0) and exercise (yes=1, no=0). Individual patient identification number was used as the frailty information. To check the proportionality (PH) assumption, ln(-ln(survival)) plot was obtained for all the binary variables. The graphs are shown in Figure 1. PH assumption was violated for alcohol intake, doing exercise and family history as the lines cross each other for each of these covariates, but there is no evidence of violation for salt intake.

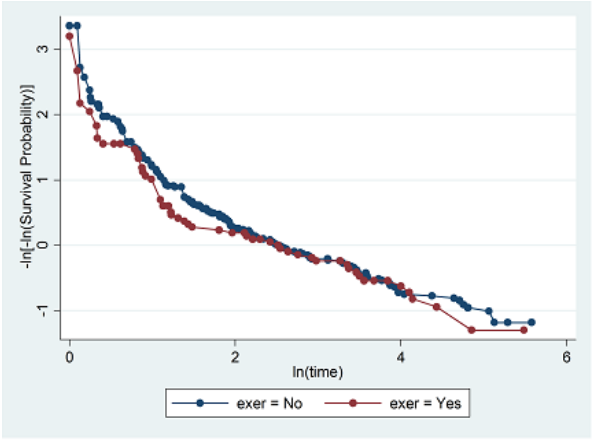
Salt intake



Alcohol intake



Exercise



Family History

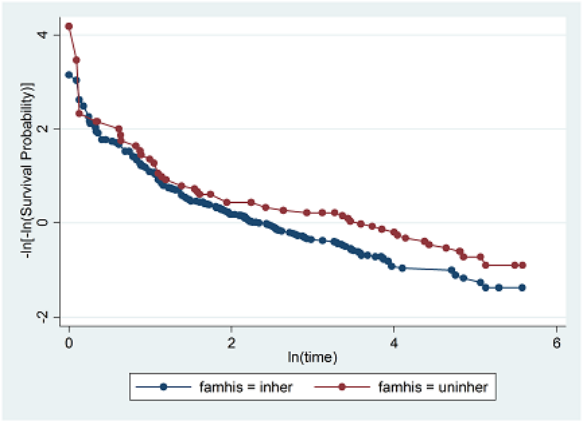


Figure 1. Testing PH assumptions on Salt intake, Alcohol intake, Doing exercise and Family history

The change point t_a as described by (4) was located at the median survival time which is 42 weeks so that the time varying function was given as

$$r(t) = \begin{cases} 1 \; if \;\; t \geq 42 \\ 0 \; if \;\; t < 42 \end{cases} \tag{15}$$

The time-varying covariate $X_i(t)$ was thus built with those covariates that violated proportionality assumption by their interactions with function of time $r(t)$) as given in (16). Non-proportional Hazards model (5) was then fitted with the two estimators PL and PPL with the four configurations PL-GAF, PL-IGF, PPL-GAF and PPL-IGF earlier described.

The results are presented in Table1. It is observed that the estimated hazard ratios (exp(β)) are similar for the two estimators under the two frailty distributions, with slightly low standard errors in favour of the penalized partial likelihood estimator. The significant variables in predicting the risk of death from diabetes are age of the patients, family history, exercise and body mass index. Older patients tend to have higher rate of mortality from diabetes. It is also noted that body mass

index is associated with increased risk of death from diabetes. As expected, those who do exercise are less likely to die from diabetes and family history is positively associated with mortality from diabetes.

Table 1. Estimates of Hazard Ratio (HR), standard error (SE) and P-value for PL and PPL estimators of diabetes data under Gamma Frailty (GAF) and Inverse Gaussian Frailty (IGF) Models.

Covariate	**PL-GAF**	**PPL-GAF**	**PL-IGF**	**PPL-IGF**
	HR (SE) **P-value**	**HR (SE)** **P-value**	**HR (SE)** **P-value**	**HR (SE)** **P-value**
Age	1.213 (0.128) 0.003	1.211(0.125) 0.007	1.217 (0.131) 0.003	1.213 (0.127) 0.002
Sex	0.875 (0.078) 0.130	0.816 (0.069) 0.242	0.872 (0.073) 0.156	0.814 (0.064) 0.137
Body Mass Index	1.073 (0.115) 0.006	1.043 (0.112) 0.025	1.068 (0.114) 0.004	1.043 (0.113 0.025)
Salt Intake	0.916 (0.108) 0.271	1.016 (0.112) 0.130	0.909(0.101) 0.210	1.101 (0.110) 0.132
Famhis	1.152 (0.121) 0.002	1.154 (0.123) 0.004	1.153 (0.123) 0.002	1.154 (0.122) 0.001
Exercise	0.688 (0.067) 0.030	0.597 (0.047) 0.031	0.686(0.068) 0.002	0.607 (0.046) 0.021
Alcohol	1.306 (0.137) 0.086	1.308 (0.144) 0.107	1.307 (0.141) 0.261	1.309 (0.146) 0.244
Famhis*Time	2.242 (0.216) <0.0001	2.235 (0.202) <0.0001	2.244 (0.220) <0.0001	2.233 (0.192) <0.0001
Exercise*Time	0.806 (0.074) <0.0001	0.743 (0.070) <0.0001	0.802 (0.071) <0.0001	0.804 (0.071) <0.0001
Alcohol*Time	1.689 (0.168) <0.0001	1.656 (0.160) <0.0001	1.701(0.172) <0.0001	1.654 (0.158) <0.0001
$\hat{\theta}$	0.503	0.508	0.442	0.439

The estimated frailty variance $\hat{\theta}$ are similar under the two estimators, with higher variances for the Gamma frailty distribution compared to Inverse Gaussian distribution. As observed, the estimates of the frailty variance ($\hat{\theta}$) for PL-GAF is 0.503 and 0.508 for PPL-GAF, whereas under Inverse Gaussian frailty, $\hat{\theta}$ for PL-IGF and PPL-IGF are 0.442 and 0.439 respectively. This implies that including gamma frailty accounts for more subject-specific heterogeneity in the data than the inclusion of Inverse Gaussian frailty.

To evaluate the effect of time - varying covariates, four progressive models, M_1, M_2, M_3 and M_4 were fitted to the data. Model comparison was done using Akaike Information Criterion (AIC) and the results are shown in Table 2.

Table 2. AIC for model comparison

Model	**Description**	**GAF**	**IGF**
M1	Model with PL and all covariates as time fixed	2503.7	2842.9
M2	Model with PPL and all covariates as time fixed	2316.8	2643.5
M3	Model with PL incorporating time varying covariates	2101.2	2209.6
M4	Model with PPL incorporating time varying covariates	2024.6	2184.1

It is observed from the table that models M3 and M4, which included alcohol intake, exercise and family history as time-varying covariates provide better fits than M1 and M2 which violated proportionality assumption on these covariates. However, penalized partial likelihood model with time-varying covariates (M4) performed the overall best with AIC of 2024.6.

In the next analysis, simulation studies were conducted to compare PL-GAF, PL-IGF, PPL-GAF and PPL-IGF configurations under the violation of proportionality assumption.

Simulation Study

Simulation studies are useful for assessing the behaviours of analytic techniques under various conditions that present complexities in practice (Montez-Rath, *et al*. (2017). Two simulations studies were conducted to evaluate the performances of the four model configurations under various scenarios. The main objective of the simulations was not to estimate the frailty variance θ but rather, to assess the performances of the PL and PPL estimators with the four configurations under study; and evaluate their sensitivity to frailty misspecification.

Simulation 1

Survival time data were simulated using model (5) with Gamma and inverse Gaussian frailty distributions. Frailty variance θ was set as 2.5. The baseline hazard function was generated from Weibull distribution to simulate Cox-weibull survival time. Censoring times were generated from exponential distribution exp(λ) at 80th, 50th and 20th percentiles of the event times distributions conditional on the covariates and frailty term to achieve the censoring proportion 20%, 50% and 80% respectively. Time invariant covariate Z_i was generated from $N(0,1)$ distribution. The time-varying covariate $X_i(t)$ was generated as an interaction of X_i with $g(t)$, where X_i was generated from a Bernoulli distribution with success probability 0.5 and $g(t)$ is a dichotomous change-point function as given in (4). The corresponding true parameters were chosen as $\gamma = -0.5$ and $\beta_t = 1$. Survival times were then simulated using the general survival simulation algorithm as in Austin (2012). For each setting, 1000 data sets were simulated with sample sizes 50, 200, 500 and 1000.

Models were then fitted with the four model configurations PL-GAF), PL-IGF), PPL-GAF and PPL-IGF.

Data Generating Procedure

The baseline hazard function for Weibull distribution with scale parameter λ and shape parameter ρ can be written as $\psi_0(t) = \lambda\rho t^{\rho-1}$. The hazard increases with time if $\rho > 0$, decreases if $\rho < 0$ and constant if $\rho = 1$. Thus the Non-proportional frailty model, with time-varying covariates (dropping subscript *i* henceforth for clarity) is given by

$$\psi(t|Z(t),W) = W\lambda p t^{p-1}\exp(\gamma'Z + \beta'X(t) + V), \tag{16}$$

with cumulative hazard function given by

$$\Lambda(t|Z(t),W) = \begin{cases} \lambda\exp(\gamma'Z+V)t^{\rho} & if\ t < t_a \\ \lambda\exp(\gamma'Z+V)[\frac{1}{\rho}(t_a^{\rho} + t^{\rho}\exp(\beta_t) - t_a^{\rho}\exp(\beta_t))] & if\ t \geq t_a \end{cases}, \tag{17}$$

and the inverse of the cumulative hazard function is given as

$$\Lambda^{-1}(t|Z(t),W) = \begin{cases} \dfrac{t}{\lambda\exp(\gamma'Z+V)} & if\ t < \lambda\exp(\gamma'Z+V)t_a \\ t - \lambda t_a[\exp(\gamma'Z+V) + \exp(\gamma'Z+\beta_t+V)] & if\ t \geq \lambda\exp(\gamma'Z+V)t_a \end{cases} \tag{18}$$

Therefore, the Cox-Weibull survival time, following Austin (2012) can be simulated as

$$T = \begin{cases} \dfrac{-\log(u)}{\lambda\exp(\gamma'Z+\beta_t+V)} & if\ -\log(u) < \lambda\exp(\gamma'Z+V)t_a \\ \dfrac{-\log(u) - \lambda t_a[\exp(\gamma'Z+V) + \exp(\gamma'Z+\beta_t+V)]}{\lambda\exp(\gamma'Z+\beta_t+V)} & if\ -log(u) \geq \lambda\exp(\gamma'Z+V)t_a \end{cases} \tag{19}$$

Comparison of the estimators were done using Relative Bias (RB) expressed in percentage and Root means Square Errors (RMSE), which are respectively given as

$$RB(\hat{\beta}_t) = \frac{1}{1000}\sum_{j=1}^{100}\left(\frac{\hat{\beta}_{tj} - \beta_t}{\beta_t}\right)\%$$

and

$$RMSE(\hat{\beta}_t) = \left[\frac{1}{1000}\sum_{j=1}^{1000}(\hat{\beta}_{tj} - \beta)^2\right]^{1/2},$$

where $\hat{\beta}_{tj}$ is the estimated time-varying parameter of the *j*th sample and β_t is the true value of the time varying effect. Lower value of RB implies a stronger agreement between the true and the estimated parameter values and this is a measure of consistency of the estimator whereas lower RMSE implies higher efficiency of the estimator. The Relative Bias (RB) and Root Mean Square Error (RMSE) of the various configurations PL-GAF, PL-IGF, PPL-GAF and PPL-IGF) for different sample sizes and censoring proportions are presented in Table 3. It is observed that the relative bias and the root mean square error decrease with increase in sample size across all scenarios. Comparing PL and PPL estimators, it is observed that RB and RMSE are consistently smaller for configuration involving PPL than for PL across all sample sizes. The proportion of censored observations also has an impact on the predictive accuracy of the estimators. As observed, for all sample sizes, both estimators become worse in terms of RB and are less efficient (higher RMSE) as the proportion of censored observations increases. One interesting finding from the study is the relatively high RMSE which results in substantial loss of efficiency for smaller sample size under both estimators when at least 50% of the observations were censored, compared to 20% censoring. For example, PL-GAF at sample size 50 has RMSE of 0.109 when 20% of observations were censored, compared to 0.356 (about 3 times higher) when 50% were censored and 0.657 (about 6 times higher) for 80% censoring. The corresponding RMSE for PPL-GAF is 0.101 when 20% of observations were censored compared to 0.334 (about 3 times higher) for 50% censoring and 0.490 (about 5 times higher) for 80 % censoring. These differences are seen to reduce substantially as the sample size increases and almost vanish at sample size 1000.

Table 3. Relative Bias (RB) and Root Mean Square Error (RMSE) for PL and PPL estimates of $\beta_t = 1.0$ under, Gamma Frailty and Inverse Gaussian Frailty distributions.

Sample Size	**% cens**	**Gamma Frailty**				**Inverse Gaussian Frailty**			
		PL-GAF		**PPL-GAF**		**PL-IGF**		**PPL-IGF**	
		PRB	**RMSE**	**PRB**	**RMSE**	**PRB**	**RMSE**	**PRB**	**RMSE**
50	20	12.7	0.109	12.2	0.101	13.2	0.119	12.6	0.106
	50	13.3	0.356	12.8	0.334	13.8	0.389	13.1	0.336
	80	13.6	0.657	13.1	0.490	14.2	0.681	13.5	0.503
200	20	12.0	0.055	11.4	0.039	12.4	0.060	11.7	0.042
	50	12.6	0.060	12.0	0.054	13.1	0.066	12.3	0.060
	80	13.0	0.094	12.4	0.080	13.5	0.103	12.7	0.088
500	20	11.7	0.047	11.5	0.046	12.1	0.051	11.6	0.046
	50	12.2	0.052	12.0	0.050	12.7	0.056	12.2	0.054
	80	12.5	0.077	12.3	0.065	13.0	0.084	12.7	0.071
1000	20	11.6	0.044	11.4	0.038	12.0	0.048	11.5	0.042
	50	12.1	0.046	11.6	0.041	12.5	0.051	11.7	0.042
	80	12.4	0.048	12.2	0.047	12.8	0.056	12.4	0.049

Simulation 2

In this simulation, sensitivity analysis was carried out to evaluate the performance of the estimators under frailty misspecification. The data generating process was as in simulation 1 but the frailty term was generated from log-skewed normal distribution given in (10). Skewness parameter was set as α = 0, 1.0, 2.5 and 5.0. The model given in (5) was then fitted by specifying Gamma and Inverse Gaussian frailty distributions with the four configurations as earlier described. Empirical power was calculated for each configuration as the proportion of cases, based on 1000 runs in which the hypothesis $H_o : \beta_t = 0$ against a two-sided alternative $H_1 : \beta_t = 1.0$ were rejected, using a Wald test at 5 % significance level. The results are shown in Figures 2 and 3. Generally, it is observed from all the configurations that frailty misspecification has an effect on the power of the estimators. From Figure 2, it is observed that the power for rejecting $H_o : \beta_t = 0$ when $\beta_t = 1.0$ generally increases with increase in sample size. Compared to PL-GAF and PL-IGF configurations, PPL-GAF and PPL-IGF configuration have higher power of rejecting H_o for all sample sizes. However, the power of PPL-IGF configuration is the highest for all scenarios considered. The power also increases as the level of skewness of the correct frailty distribution (log-skewed normal) increases in all the configurations. Also, as observed from Figure 2, the superiority of PPL-GAF and PPL-IGF over PL-GAF and PL-IGF is higher in terms of power at sample size 50 and 200, but get close at sample sizes 500 and 1000.

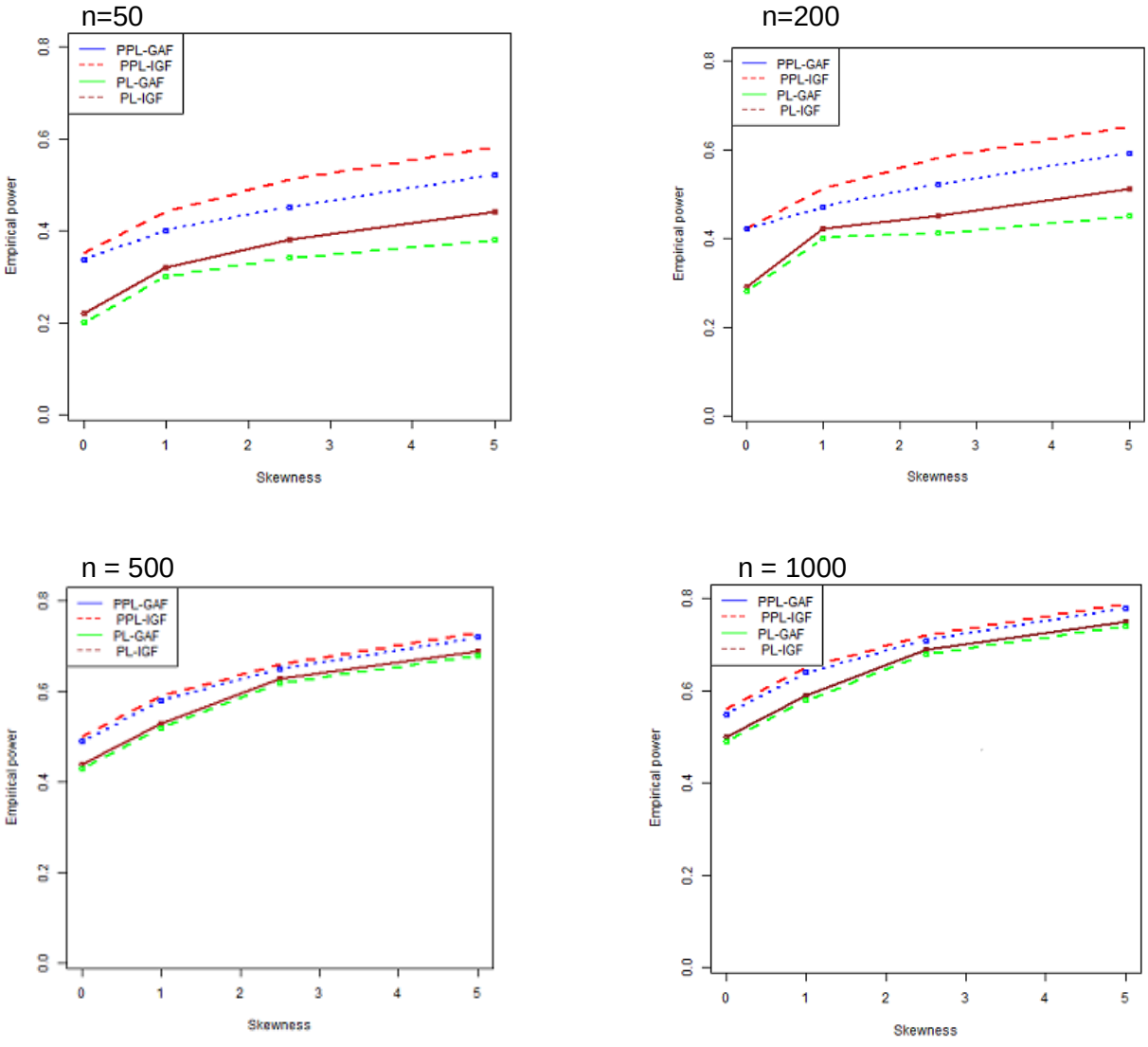


Figure 2. Empirical power for rejecting the hypothesis $H_o : \beta_t = 0$ against $H_1 : \beta_t = 1.0$ assuming a Gamma and Inverse Gaussian frailty distributions when the true distribution is log skewed normal with skewness parameters 0, 1.0, 2.5, 5.0 and sample sizes 50, 200,500,1000

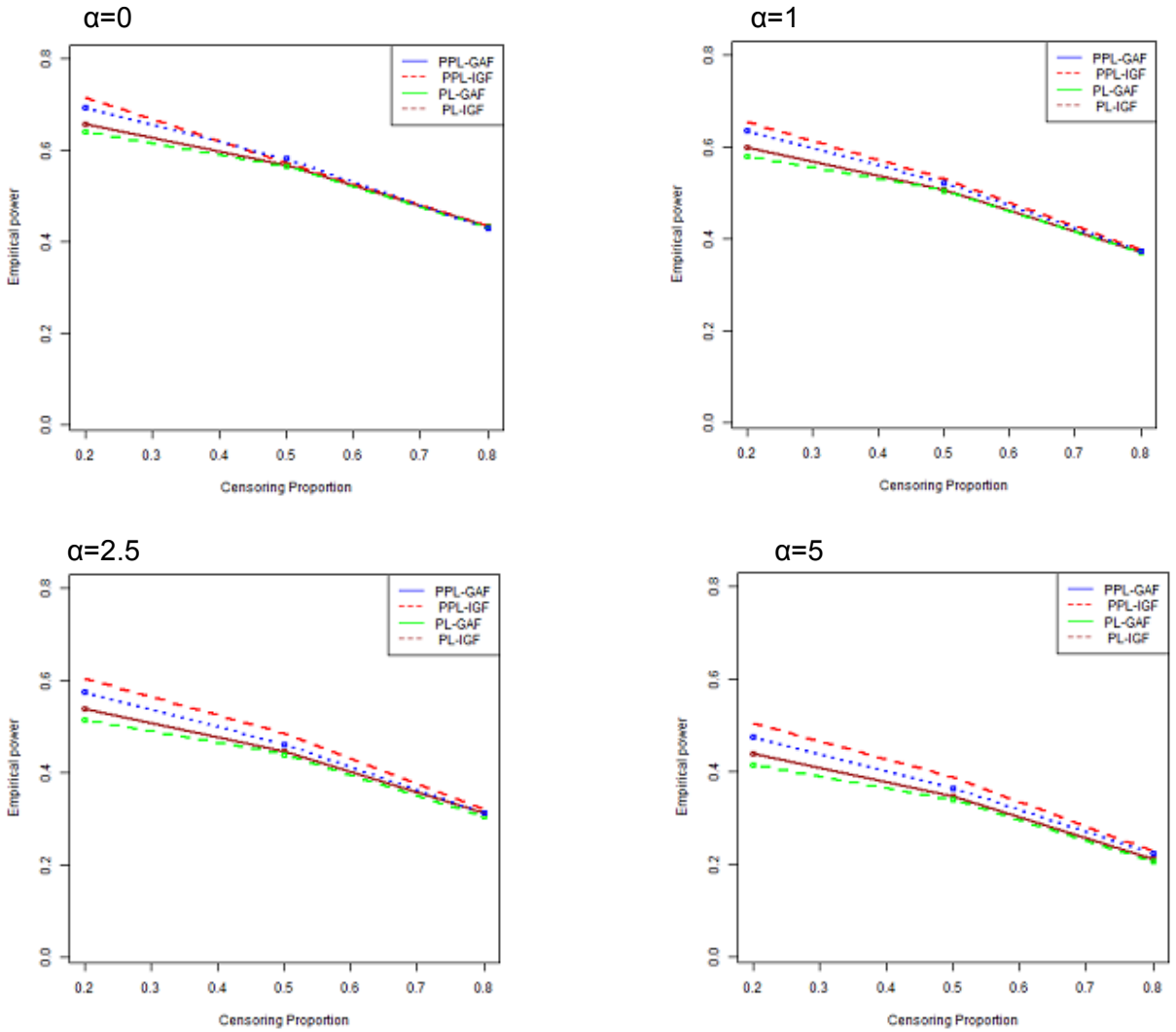


Figure 3. Empirical power for rejecting the hypothesis $H_o : \beta_t = 0$ against $H_1 : \beta_t = 1.0$ assuming a Gamma and Inverse Gaussian frailty distributions when the true distribution is log skewed normal with skewness parameters 0, 1.0, 2.5, 5.0. and censoring percentage 20%, 50% ,80%, n=200

In Figure 3, it is clearly observed that empirical power decreases as the proportion of censored observations increases for all skewness levels. It is also observed that the PPL-GAF and PPL-IGF are more powerful than the corresponding PL-GAF and PL-IGF and their power advantage are more pronounced at censoring proportion 20%. This advantage is seen to reduce as the level of censoring increases and converge at 50% or higher censoring levels. The convergence in power of the four configurations is sharper at the skewness levels α=0 and α=1 than at α=2.5 and α=5.

Discussion

The main purpose of this study was to examine the performances of two estimators, namely the partial likelihood (PL) estimator and penalized partial likelihood (PPL) estimator in Non-proportional survival models with subject-specific frailty. A dataset on diabetes was analysed and it was established under various scenarios that PPL estimator that treated the frailty densities in the complete data likelihood as penalty term outperformed PL estimator. Simulation studies were conducted to compare the two estimators. Efficiency of the two estimators was measured by the root mean square error (RMSE) and the consistency (agreement between the true and estimated parameters) measured by relative bias (RB). Results from simulation 1 showed from different likelihood-frailty configurations, namely PL-GAF, PL-IGF, PPL-GAF and PPL-IGF, that PPL estimator generally had better performance than PL estimator in terms of efficiency and it was less biased. It was also found, as expected that efficiency increased with increase in sample size, which was a satisfactory performance. Censoring proportion was negatively correlated with the performance of both estimators. From the sensitivity analysis conducted to assess the effect of frailty misspecification on the estimators, it was found that sample size, proportion of censored observation and the shape of the frailty distribution (log-skewed) severely affected the power of the estimators. It was also revealed that configurations with PPL were better in performances than their PL counterparts under all scenarios considered in this study.

Conclusion

It was reported in Neuhaus et al.(1992) that one cannot establish effects of misspecification through fitting models on the basis of a single random effects distribution. In our current study, findings from fitting two random effects (frailty) distributions, the Gamma and Inverse Gaussian distributions revealed that misspecification tends to reduce the power of estimator, and this confirmed the claims of Litière et al. (2007), which states that random effects misspecification can produce marked decreases in power and also a similar study by Heckrnan and Singer (1984), revealed high sensitivity of parameter estimates to the assumed functional form for the distribution of unobserved variables. Our study has therefore proven that penalized partial likelihood method of estimation is a preferred candidate to partial likelihood method in the analysis involving Non-proportional hazards models, in the presence of frailty and its misspecification.

References

Adeleke, K.A., Abiodun, A. A. & Ipinyomi, R. A. (2015). Semi-Parametric Non-Proportional Hazard Model with Time Varying Covariate. Journal of Modern Applied Statistical, Methods, 14 (2), 68-87.

Andersen, P. K., Klein, J. P., Knudsen, K. & Palacios, R. T. (1997). Estimation of variance in Cox's regression model with shared gamma frailties. Biometrics, 53, 1475–1484.

Austin, P.C., Mamdani, M.M., van Walraven C. & Tu, JV. (2006). Quantifying the impact of survivor treatment bias in observational studies. Journal of Evaluation in Clinical Practice, 12(6), 601–612.

Austin P.C. (2012). Generating survival times to simulate Cox proportional hazards models with time-varying covariates. Statistics in Medicine, 31, 3946–3958.

Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. Scandinavian Journal of Statistics, 12, 171-178.

Beyersmann, J., Wolkewitz, M. & Schumacher M. (2008). The impact of time-dependent bias in proportional hazards modelling. Statistics in Medicine, 27, 6439–6454.

Cox, D.R. (1972). Regression Models and Life-tables (with discussion). J. Roy. Statist. Soc. Ser. B. 34, 187-220.

Dechateau, L. & Janssen, P.(2004). Penalized Partial Likelihood for Frailties and Smoothing Splines in Time to First Insemination Models for Dairy Cows. Biometrics, 60 (3): 608-614.

Duchateau, L. & Janssen, P. (2008). The frailty model, Springer-Verlag, New York.

Heckman, J. & Singer, B. (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. Econometrica, 52 (2), 271-320

Huang, J. & Harrington , D. (2002). Penalized Partial Regression for Right-censored Data with Bootstrap Selection of the Penalty Parameter. Biometrics, 58, 781- 791.

Hougaard, P. (2000). Analysis of multivariate survival data. Springer, New York 213, 701– 714.

Litière, S. Alonso, A. & Molenberghs, G. (2007). Type I and type II error under random effects misspecification in generalized linear mixed models. Biometrics. 63:1038–1044. [PubMed: 17425642]

Lee, S., Seo, M. H. and Shin, Y. (2011). Testing for threshold effects in regression models. J. Amer. Statist. Assoc. 106, 220-231.

Liwen, L., Wang, H. J. & Zhu, Z. (2014). Testing for Change Points due to a Covariate Threshold in Quantile Regression. Statistica Sinica 24 (2014), 1859-1877 doi:http://dx.doi.org/10.5705/ss.2012.322

McGilchrist, C. A. & Aisbett, C.W. (1991). Regression with frailty in survival analysis, Biometrics, 47, 461-466.

McGilchrist, C. A. (1993). REML estimation for survival models with frailty. Biometrics, 49, 221–225

Montez-Rath, M., Kapphahn, K., Mathur, M.B., Mitani, A. A., Hendry, D. J. & Desai, M. (2017). Guidelines for Generating Right-Censored Outcomes from a Cox Model Extended to Accommodate Time-Varying Covariates. Journal of Modern Applied Statistical Methods, 16 (1), 86-106. doi: 10.22237/jmasm/1493597100

Munda, M., Rotolo, F. & Legrand. C. (2012). parfm: Parametric Frailty Model in R. Journal of Statistical Software, 15(11), 1-20.

Neuhaus, J.M, Hauck, W.W. & Kalbfleisch, J.D. (1992). The effects of mixture distribution misspecification when fitting mixed-effects logistic models. Biometrika 79:755–762.

Ng, S.K. & McLachlan, G. J. (2003). An EM-based semi-parametric mixture model approach to the regression analysis of competing-risks data. Statistics in Medicin, 22:1097 –1111

Nielsen, G. G., Gill, R. D., Andersen, P. K. & Sørensen, T. I. A. (1992). A counting process Approach to maximum likelihood estimation in frailty models. Scand. J. Statist, 19, 25–44.

Pickles, A., & C'rouchley. R. (1995). A Cornparison of Frailty Models for Multivariate Survival Data. Statistics in Medicine. 14, 1447-1461.

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

Suissa, S. (2007). Immortal time bias in pharmacoepidemiology. American Journal of Epidemiology, 167(4), 492–499.

Therneau, T. M. & Grambsch, P. M. (2000). Modeling Survival Data. Extending the Cox Model. New York. Springer Verlag.

Therneau, T. M., Grambsch, P. M., & Pankratz, V. S.(2003). Penalized survival models and frailty. Journal of Computational and Graphical Statistics, 12, 156–175.

Vaida, F. & Xu, R. (2000). Proportional hazards model with random effects. Stat. Med, 19, 3309–3324.

Valenta, Z. & Weissfeld, L. (2002). Estimation of the survival function for Gray's piecewise time-varying coeffients model. Statistics in Medicine, 21, 717–727.

Vaupel, J.W., Manton, K.G. & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography, 16, 349-454.

Verweij, P. J. M. & Houwelingen, H.C.V (1993). Cross-validation in survival analysis. Statistics in Medicine, 12 (24), 2305–2314.

Wienke, A. (2010). Frailty Models in Survival Analysis. Chapman & Hall/CRC, Boca Raton.

Zhou M. (2001). Understanding the Cox regression models with time-change covariates. The American Statistician; 55(2):153{155.DOI:10.1198/000313001750358491.

How to cite this paper:
Alfred, A. A. , Benjamin, A. O. & Kazeem, A. A. (2019). An Empirical Comparison between Partial Likelihood and Penalized Partial Likelihood Estimators for Semi-parametric Non-proportional Hazards Models with Frailty. *Malaysian Journal of Applied Sciences, 4*(1), 1-14.