African Journal of Applied Statistics
Vol. 6 (1), 2019, pages 519 –535.
DOI: http://dx.doi.org/10.16929/ajas/2019.519.228

AJAS / SPAS
ISSN 2316-0861

Modelling Clustered Survival Data with Competing Risks

Olubimpe Mercy Oladuti[1,2,*] **, and Benjamin Agboola Oyejola**[2]

[1]Department of Statistics, Federal University of Technology, Akure
[2]Department of Statistics, University of Ilorin, Ilorin

Received on August 17, 2018; Published online March 05, 2019

Abstract. . The classical Cox proportional hazards model is the most popular semiparametric model commonly used to assess the effects of risk factors in a homogeneous population for continuous survival time data. It is however based on the assumption that survival times are untied. Discrete-time logit model has been widely applied to survival time data in order to handle ties when there is a single cause for the event occurring. There are situations when an individual is at risk of experiencing several risks of failure and only the first of them is observed. In this study, binary logit model for competing risk with mixture of baseline hazard functions for clustered discrete-time data is proposed.

Key words: Cox model; Discrete-time logit model; Competing risks; Clustered; Simulation Study
AMS 2010 Mathematics Subject Classification Objects : 62Nxx;62J12

Presented by Dr. Diam Ba, Affiliated to University Gaston Berger, Senegal
Corresponding Member of the Editorial Board.

* Olubimpe Mercy Oladuti: omoladuti@futa.edu.ng
Benjamin Agboola Oyejola: boyejola2003@yahoo.com

Abstract (French) La méthode des risques proportionnels de Cox est la méthode semi-paramétrique la plus populaire qui est utilisée pour établir les risques dûs aux facteurs en jeu dans une population homogène dans l'analyse de survie avec un temps continu. Cependant, cette méthode repose sur l'hypothèse que les temps de survie ne peuvent se répéter. La méthode logit basée sur un temps discret a été largement appliquée pour gérer les égalités de temps de survie s'il y a une seule cause génératrice des évènements. Dans certaines situations, un individu peut être à risque selon plusieurs facteurs et souvent seule la survenance du premier échec est pris en compte. Dans ce papier, le modèle logit binaire pour des risques compétitifs, reposant sur des mélanges des fonction de risques pour des données en grappes est étudié selon un temps discret.

1. Introduction

Survival analysis is the analysis of data measured from a specific time of origin until an event of interest or a definite endpoint Collect (1982). The time of interest is usually characterized by means of the hazard function, signifying the rate of occurrence of the event at a given time t, but mostly via the survival function, representing the probability of *surviving* up to time t. That is, the probability that the event has not yet occurred before time t. In the original setting of survival analysis, there is a single cause for the event to occur but there are situations where several causes of failure are possible; only the occurrence of the first of them can however be observed provided only one cause is of interest. This situation is known as competing risks. This is because the smallest realized time, the cause specific failure time, makes the failure times for other causes right censored. That is, the minimum of the failure times is only observed. Clustered survival time data are commonly encountered in scientific investigations where each study subject may experience several types of event or when there are clustering of observational units such that failure times within the same cluster are correlated. According to Hougaard (2000), clustered survival analysis generally deals with survival times of multiple individuals whose failures can be dependent, repeated occurrences of the same event, known as multiple data and times to several events an individual may experience known as multiple events. Competing risks analysis addresses a type of clustered survival data that is definitely different from the types of data which Hougaard (2000) used.

The clustered survival data concentrate on estimating the parameters of one type of event at a time when outcomes are independent based on the theory that the censoring mechanism is independent of the event type of interest. This assumption is debased when multiple types of events occur but only the occurrence of the first of them can be observed.Nevertheless, in a situation where outcomes or the failure times are not all independent or when units cluster and more than two outcomes are to be observed per cluster. The clustered modelling is used.

Cox proportional hazard model is one of the common methods for analyzing survival time data. It is a robust model and its results were closely approximate the results for the correct parametric model (See Kleinbaum and Klein (2012)). It is the most common semi-parametric model used to evaluate the effects of risk factors in a population for continuous survival time data under the basic assumption that survival times are untied Anderson and Fleming (1995). In practice there is always some smallest time unit that ties can occur especially when two or more individuals experience events of interest simultaneously. In this case, continuous time model is very sensitive(See Allison (1982), Allison (2010), Singer and Willet (1993), Singer and Willet (1995), Singer and Willet (2003) and Hosmer and Lemeshow (1999)).

A discrete-time survival modeling for discrete-time data was then proposed by Cox (1972), Allison (1982), Allison (2010), Singer and Willet (1993), Singer and Willet (1995), Singer and Willet (2003) to handle ties. In trial settings, the discrete-time survival model is use for longitudinal studies when the data are often collected at discrete-time periods. It examines the shape of hazards function, and it is simple to execute using the logistic regression model (see Xie *et al.*(2003), McCallon (2009), Sharaf and Tsokos(2014). Cox (1972) introduced the discrete-time hazard model in terms of logit-hazard rather than hazard in his article and have been in use for decades (see Allison (1982), Allison (2010),Willet and Singer (1991), Willet and Singer (1993), Singer and Willet (1993), Singer and Willet (1995), Singer and Willet (2003)) but they are less visible than continuous time survival model, especially in the medical and behavioral sciences area (See Altman *et al.*(1995), Enderlein *et al.* (1986), Barber *et al.* (2000), Xie *et al.*(2003), McCallon (2009)).

The discrete-time survival model have been in use for decades, but they are less visible than continuous time survival model, especially in the medical and behavioral sciences area (See Altman *et al.*(1995), Enderlein *et al.* (1986)). The discrete-time survival model was proposed by Cox (1972), and it is a type of logistic regression. The discrete-time models often used for survival analysis are logit, probit and complementary log-log. The analysis for this type of model needs a properly structured data set with multiple records per subject. The discrete-time model are used more appropriately in the situation with large number of ties when more than two individuals experience an event at the same time (see Allison (1982), Allison (2010)) when the time are truly discrete and when the time of experiencing event is hard to tell. Most studies handle ties when there is single cause for the event to occur. But there are situations where an individual can experience several causes of failure (competing Risk) and more than one individual experience their first event at the same time. Due to this, there is a need to manage the ties with discrete-time competing risk model using a person-period format in order to avoid biased estimates.
,

In this study, a clustered discrete-time binary logit model under competing risk setting is proposed and then compared with Cox model under the mixture of baseline hazard distributions marginally and jointly.

2. Theoretical background

2.1. Competing Risk Formulation with Cox model

Consider *M* independent clusters ($m = 1, \ldots, M$), and there are n_m individuals (subjects) in cluster *m*. Suppose that for *j*th individual (subject), G types of failure may occur. Let $T_{ij}^g, C_{ij}^g\ and\ X_{ij}^g$ be the independent failure time, censoring time and p-vector of possible covariates respectively for *j*th individual in *i*th cluster experiencing g^{th} type of failure($i = 1, \ldots, m\ ;; j = 1, \ldots, n_m, g = 1, \ldots, G$). Let $T_{ij} = t_{ij}$ and C_{ij} be the time to failure and the censoring time for j^{th} individual in i^{th} cluster respectively and x_{ij} be a vector of covariates. Assume that T_{ij} and C_{ij} are independent conditional on the covariate vector, X_{ij}. We define $T_{ij} = min\left(t_{ij}, c_{ij}\right)$ and $d_{ij} = I\left(t_{ij} \leq c_{ij}\right)$ where $I(.)$ is an indicator function which indicates whether or not the main event of interest has occurred; it is equal to one if the condition is true and zero otherwise.

In survival modeling, an appropriate method must be chosen to handle the different event types when both events are of interest. In the competing risks approach, a separate model is specified for the timing of each type of event and each of these models can be estimated separately for single event. The proportional hazards model with competing risk for j^{th}individual in i^{th} cluster can be written as:

$$\xi\left(t_{ij}^g/x_{ij}\right) = \xi_o\left(t_{ij}\right) exp(\beta^{g'} x_{ij}^g) \tag{1}$$

where $\xi(t_{ij}^g/x_{ij})$ is the hazard at time **t** for j^{th} individual in i^{th} cluster having event type g with covariate value x_{ij}^g, $\xi_o(t_{ij})$ is the baseline hazard at time **t**, ${}^{g'}x_{ij}^g$ is the effect of the covariate on the hazard for event type **g**.

2.2. Proposed Discrete-Time Logit Model with Competing Risk

Suppose that the timeline for individual is partitioned into *l* mutually exclusive intervals $[0, a_1), [a_1, a_2), [a_2, a_3), [a_3, a_4), \ldots, [a_{t-1}, a_t), [a_t, a_\infty)$ in each cluster so that we observe discrete time $T \in 1, \ldots, t$ where *T* = *t* denotes failure within the interval$[a_{i-1}, a_i]$. In discrete time, for each time interval t, a vector of binary response is defined as $y_{t_{ij}}^g = \left(y_{t_{ij}}^1, y_{t_{ij}}^2, \ldots, y_{t_{ij}}^k\right)$, where t_{ij} is the observed time in the interval for which individual j in the cluster i is observed and a binary response $y_{t_{ij}}^g$ is created for each event time interval t up to c_{ij} which is coded as follows

$$y^g_{t_{ij}} = \begin{cases} 0 & t < c_{ij} \\ 0 & t = c_{ij}, \quad d_{ij} = g \\ 1 & t = c_{ij}, d_{ij} \neq g \end{cases} \tag{2}$$

All individuals regardless of whether or not their duration is censored will have $y^g_{t_{ij}} = 0$ for interval $t = c_{ij}$.Then the time $t^g_{ij}\,(g = 1, \ldots k)$ is the time at which event type g occur to an individual j in the cluster/group i, for uncensored individual $t^g_{ij} = min\left(t^1_{ij}, \ldots, t^k_{ij}\right)$.

The failure process of individual i with failure event type g can then be considered as a sequence of binary response outcomes which follow a binomial distribution. The binary event indicator can then be define as:

$$y^g_{t_{ij}} = \begin{cases} 1 & if\ t = t^g_{ij}\ and\ d_{ij} = 1 \\ 0 & Otherwise \end{cases} \tag{3}$$

where $d_{ij}$is the censoring indicator which takes value 1 if individual j in cluster i has failure event type g at time t and value 0 if otherwise. Let $\frac{pr(y^g_{tij}=1)}{pr(y^g_{tij}=0)}$ be the odds of event type g occurring in interval $[a_{t-1}, a_t)$.

Cox (1972) proposed an extension of the proportional hazards model to discrete time by working with the conditional odds of an event of failure occurring at each time t_{ij} given survival up to that point. Extending this to competing risk setting, the discrete-time for clustered survival time data for event type g can be obtained as

$$\frac{\xi^g_{tij}|x^g_{ij}}{1-\xi^g_{tij}|x^g_{ij}} = e^{(a^g_{ij}+\beta^g_i x^g_{ij})}$$

and

$$\xi(t^g_{ij}/x_{ij}) = \frac{e^{(a^g_{ij}+\beta^g_i x^g_{ij})}}{1 + e^{(a^g_{ij}+\beta^g_i x^g_{ij})}} \tag{4}$$

$$1 - \xi(t^g_{ij}/x_{ij}) = \frac{1}{1 + e^{(a^g_{ij}+\beta^g_i x^g_{ij})}} \tag{5}$$

$$\frac{\xi(t^g_{ij}/x_{ij})}{1 - \xi(t^g_{ij}/x_{ij})} = e^{(a^g_{ij}+\beta^g_i x^g_{ij})}$$

$$\frac{\xi(t^g_{ij}/x_{ij})}{1 - \xi(t^g_{ij}/x_{ij})} = \frac{\xi_o(t_{ij})}{1 - \xi_o(t_{ij})} exp\left\{\beta^{g'} x^g_{ij}\right\} \tag{6}$$

$$Logit\left[\frac{\xi(t^g_{ij}/x_{ij})}{1 - \xi(t^g_{ij}/x_{ij})}\right] = a^g_{ij} + \beta^{g'} x^g_{ij} \tag{7}$$

where $\xi(t_{ij}^g/x_{ij})$ is the hazard at time t for j^{th}individual in i^{th} cluster having event or failure type g with covariate value x_{ij}^g, $\xi_o(t_{ij})$ is the baseline hazard at time **t** and $\beta^{g'}x_{ij}^g$ is the relative risk associated with covariate values x_{ij}^g.

By taking the log, a model on the logit of the hazard or conditional probability of experiencing event type g at t_{ij}^g given survival up to that time is given as follows;

$$Log\left[\frac{\xi(t_{ij}^g/x_{ij})}{1-\xi(t_{ij}^g/x_{ij})}\right] = a_{ij}^g + \beta^{g'}x_{ij}^g. \tag{8}$$

where $\xi(t_{ij}^g/x_{ij})$ is the hazard at time t for j^{th}individual in i^{th} cluster having event or failure type g with covariate value x_{ij}^g, a_{ij}^g=log($\frac{\xi_o(t_{ij})}{1-\xi_o(t_{ij})}$) is the baseline effect and $\beta^{g'}$ is the relative risk associated with covariate values x_{ij}^g.

Clearly, (8) can be written as

$$logit\left[\xi(t_{ij}^g/x_{ij})\right] = a_{ij}^g + \beta^{g'}x_{ij}^g. \tag{9}$$

where a_{ij}^g=$logit[\xi_o\,(t_{ij})]$,

3. Simulation study

Survival times were generated to simulate Cox models with known regression coefficients considering the Exponential, the Weibull, the Gompertz and the Lognormal distribution as baseline hazard. The general relationship between the hazard and the corresponding survival time of the usual Cox model was developed as in Bender *et al.* (2005).

The Cox proportional hazards model is given by

$$T = H_o^{-1}\left[-\log\left(U\right) exp\left(-\beta^{g'}x_{ij}^g\right)\right]$$

i.e.

$$T = H_o^{-1}\left[\frac{-\log(\mathrm{U})}{exp\left(\beta^{g'}x_{ij}^g\right)}\right],$$

where U is the random variable with $U \sim uni(0,1)$, $\beta^{g'}x_{ij}^g$ is the effect of the covariates on the hazard for failure event type g=1,2, H_o^{-1} is the inverse of a cumulative baseline hazard function.

We assume that the baseline hazard H_o can be Weibull, Exponential, Lognormal or Gompertz distribution.

3.1. Weibull Distribution

$$T = \lambda^{-1}\left[-\log(U)\,exp\left(-\beta^{g'}x^{g}_{ij}\right)\right]^{1/v}$$

$$= \left[\frac{-\log(\mathrm{u})}{\lambda exp\left(\beta^{g'}x^{g}_{ij}\right)}\right]^{1/v}, \qquad \lambda > 0,\ v > 0,$$

with baseline hazard function $h_o(t) = \lambda v t^{v-1}$ where λ is scale parameter and v is the shape parameter

3.2. Exponential Distribution

$$T = \lambda^{-1}\left[-\log(U)\,exp\left(-\beta^{g'}x^{g}_{ij}\right)\right]$$

$$= \left[\frac{-\log(\mathrm{u})}{\lambda exp\left(\beta^{g'}x^{g}_{ij}\right)}\right] \quad ,\lambda > 0,$$

with constant baseline hazard function $h_o(t) = \lambda$

3.3. Gompertz Distribution

$$T = \frac{1}{\alpha}log\left[-\frac{\alpha}{\lambda}(\log(U)\,exp\left(-\beta^{g'}x^{g}_{ij}\right)) + 1\right]$$

$$= \frac{1}{\alpha}log\left[1 - \frac{\alpha\log(\mathrm{u})}{\lambda exp\left(\beta^{g'}x^{g}_{ij}\right)}\right] \quad ,\lambda > 0\ -\infty < \alpha < \infty,$$

with baseline hazard function $h_o(t) = \lambda\exp(\alpha t)$

3.4. Lognormal Distribution

This is one of the commonly used distributions in survival time but do not have property of proportional hazards like the other parametric distributions above. David and Albert (2014) derived lognormal survival time as follows;

$T = exp\left(\mu_j + s\left(\log(\mu)\ - \log(1-u)\right)\right)$, s>0; t>0 where $\mu_j = \left(\beta^{g'}x^{g}_{ij}\right)$

with baseline hazard function $h_o(t) = (ts)^{-1}$

Data that follow various survival distributions were generated to compare the model under different scenarios. Simulation studies were carried out for two events type with mixture of the baseline hazard distributions with fixed parameters. Dataset with two covariates X_1 from a Normal N[0; 1] and X_2 from a Binomial B[1, 0.5] was generated. The corresponding true regression coefficients are fixed as β_1 = 1, β_2= -1. Sample sizes are 100, 200,500, 1000 and 2000 with a censoring rate of 35%. For each parameter combination, all simulated data was replicated 1000 times. The simulated datasets were expanded into a person-period format in order to fit the discrete time logit survival model for the models specified for each event and minimum of the two events. All of the datasets were simulated and modeled in R statistical software.

4. RESULTS

The summary of results are presented in the tables below where the mixtures of parametric baseline hazard distributions are Weibull-Lognormal (WL), Weibull-Exponential (WE), Lognormal-Exponential (LE), Weibull-Gompertz (WG), Lognormal-Gompertz (LG), Exponential-Gompertz (EG).

The results of simulation study are summarized in Tables 1-6 for the estimates of, mean absolute bias (MAB) and mean square error for prediction (MSEp) for Cox model and discrete-time logit model (DTLM) with mixture of distributions for the baseline hazards under different sample sizes.
The results in Table 1 above shows that estimated values are close to the true values. Increase in sample size decreases the estimates and mean absolute bias (MAB) for both the Cox and Discrete-time model. The results also showed that increase in sample size decreases the MSEp for both models, although the MSEp were rapidly decrease for DTLM compare to Cox model.

When considering the Overall Event, the estimated values are closer to the true values but DTLM provides more precise estimates than Cox Model. Also, Discrete-time model performs better than Cox model in terms of estimated values.
The estimate mean values from Table 2 are close to true parameter values. The estimated mean value and mean absolute bias (MAB) decreases as sample size increases but DTLM estimates are overestimated. The MSEp for both the Cox model and DTLM decrease as the sample size increases.

For the Overall Event, DTLM has more precise estimates compared to Cox model, there is no loss of efficiency in terms of MSEp. Considering the two models in term of estimated mean values, Discrete-time model performs better than Cox model.

The results in Table 3 above indicate that estimate mean values are consistently close to the true values with minimum MSEp. Also, mean absolute bias (MAB) and the MSEp are decreasing as a result of increase in sample size but DTLM gives a minimum MSEp compare to Cox model.

In view of the Overall Event, the estimated mean values are close to the true values but DTLM gives an over-estimated values. Comparing the estimates of the two models, Discrete-time model performs well than Cox model in term of estimates.
Table 4 results reveal that estimate mean values are close to the true values. Increase in sample size decreases the estimates and mean absolute bias (MAB). The Cox model are underestimated for Event2 (Lognormal) with large MSEp.

Table 1: Estimates, Absolute Bias and Mean Squared Errors for Weibull-Exponential mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$)
Legend : Event 1 (Weibull), Event 2 (Exponential)

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1	100	$\hat{\beta}_1$	0.951(0.092)	0.996(0.077)	0.013	0.009
		$\hat{\beta}_2$	-1.093(0.165)	-1.164(0.193)	0.043	0.056
	200	$\hat{\beta}_1$	0.998(0.055)	1.185(0.185)	0.005	0.039
		$\hat{\beta}_2$	-0.807(0.200)	-0.972(0.103)	0.052	0.017
	500	$\hat{\beta}_1$	1.094(0.095)	1.161(0.161)	0.012	0.028
		$\hat{\beta}_2$	-0.884(0.121)	-0.916(0.097)	0.020	0.014
	1000	$\hat{\beta}_1$	0.948(0.053)	1.074(0.074)	0.004	0.007
		$\hat{\beta}_2$	-1.021(0.048)	-1.147(0.147)	0.003	0.025
	2000	$\hat{\beta}_1$	0.915(0.085)	1.015(0.022)	0.008	0.003
		$\hat{\beta}_2$	-0.954(0.050)	-1.040(0.045)	0.003	0.003
Event 2	100	$\hat{\beta}_1$	0.960(0.090)	0.989(0.080)	0.013	0.010
		$\hat{\beta}_2$	-1.106(0.176)	-1.155(0.194)	0.049	0.055
	200	$\hat{\beta}_1$	1.002(0.055)	1.157(0.157)	0.005	0.029
		$\hat{\beta}_2$	0.820(0.189)	-0.957(0.105)	0.047	0.017
	500	$\hat{\beta}_1$	1.093(0.095)	1.132(0.132)	0.011	0.020
		$\hat{\beta}_2$	-0.867(0.135)	-0.877(0.127)	0.024	0.021
	1000	$\hat{\beta}_1$	0.943(0.058)	1.033(0.038)	0.004	0.002
		$\hat{\beta}_2$	-1.022(0.047)	-1.114(0.115)	0.003	0.016
	2000	$\hat{\beta}_1$	0.915(0.085)	0.989(0.020)	0.008	0.001
		$\hat{\beta}_2$	-0.959(0.046)	-1.024(0.037)	0.003	0.002
Overall Event	100	$\hat{\beta}_1$	0.977(0.049)	1.030(0.050)	0.004	0.004
		$\hat{\beta}_2$	-1.124(0.136)	-1.204(0.205)	0.025	0.051
	200	$\hat{\beta}_1$	1.021(0.035)	1.226(0.226)	0.002	0.053
		$\hat{\beta}_2$	-0.833(0.168)	-1.011(0.056)	0.032	0.005
	500	$\hat{\beta}_1$	1.126(0.126)	1.199(0.199)	0.017	0.040
		$\hat{\beta}_2$	-0.897(0.103)	-0.933(0.070)	0.013	0.006
	1000	$\hat{\beta}_1$	0.971(0.030)	1.100(0.100)	0.001	0.010
		$\hat{\beta}_2$	-1.048(0.049)	-1.181(0.181)	0.003	0.033
	2000	$\hat{\beta}_1$	0.940(0.060)	1.045(0.045)	0.004	0.002
		$\hat{\beta}_2$	-0.983(0.021)	-1.077(0.077)	0.001	0.006

The MSEp for both Cox model and DTLM decreases as sample size increasing, but the MSEp were higher for Cox compare to the DTLM model.

For the Overall Event, increase in sample size decreases the estimates with MAB for Cox model but has no effect for DTLM and there is no loss of efficiency in terms of MSEp. Comparing the two models, Discrete-time model performs well than Cox model in term of estimated mean values except Event2 (Lognormal) that we have under-estimated value.

Table 2: Estimates, Absolute Bias and Mean Squared Errors for Weibull-Gompertz mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$)
Event 1 (Weibull), Event 2 (Gompertz)

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1	100	$\hat{\beta}_1$	0.948(0.092)	0.992(0.077)	0.013	0.009
		$\hat{\beta}_2$	-1.097(0.166)	-1.158(0.189)	0.044	0.054
	200	$\hat{\beta}_1$	1.003(0.055)	1.183(0.183)	0.005	0.038
		$\hat{\beta}_2$	-0.810(0.197)	-0.970(0.103)	0.051	0.017
	500	$\hat{\beta}_1$	1.094(0.096)	1.157(0.157)	0.012	0.027
		$\hat{\beta}_2$	-0.882(0.124)	-0.910(0.101)	0.021	0.015
	1000	$\hat{\beta}_1$	0.950(0.052)	1.072(0.072)	0.004	0.006
		$\hat{\beta}_2$	-1.022(0.048)	-1.144(0.144)	0.004	0.024
	2000	$\hat{\beta}_1$	0.915(0.085)	1.012(0.020)	0.008	0.001
		$\hat{\beta}_2$	0.955(0.050)	-1.037(0.043)	0.003	0.003
Event 2	100	$\hat{\beta}_1$	0.957(0.091)	0.985(0.080)	0.013	0.01
		$\hat{\beta}_2$	-1.111(0.178)	-1.149(0.190)	0.05	0.053
	200	$\hat{\beta}_1$	1.007(0.055)	1.155(0.156)	0.005	0.029
		$\hat{\beta}_2$	0.823(0.186)	-1.955(0.106)	0.046	0.018
	500	$\hat{\beta}_1$	1.093(0.095)	1.124(0.124)	0.011	0.018
		$\hat{\beta}_2$	-0.861(0.141)	-0.866(0.137)	0.026	0.024
	1000	$\hat{\beta}_1$	1.944(0.057)	1.031(0.037)	0.004	0.002
		$\hat{\beta}_2$	-1.023(0.047)	-1.111(0.112)	0.030	0.015
	2000	$\hat{\beta}_1$	0.914(0.086)	0.982(0.023)	0.008	0.002
		$\hat{\beta}_2$	-0.963(0.044)	-1.023(0.036)	0.030	0.02
Overall Event	100	$\hat{\beta}_1$	0.974(0.049)	1.027(0.049)	0.004	0.004
		$\hat{\beta}_2$	-1.127(0.138)	-1.198(0.200)	0.026	0.048
	200	$\hat{\beta}_1$	1.026(0.037)	1.225(0.225)	0.002	0.052
		$\hat{\beta}_2$	-0.836(0.164)	-1.009(0.056)	0.031	0.005
	500	$\hat{\beta}_1$	1.126(0.126)	1.191(0.191)	0.017	0.037
		$\hat{\beta}_2$	-0.892(0.108)	-0.922(0.079)	0.014	0.008
	1000	$\hat{\beta}_1$	0.972(0.029)	1.099(0.098)	0.001	0.010
		$\hat{\beta}_2$	-1.049(0.050)	-1.178(0.178)	0.003	0.032
	2000	$\hat{\beta}_1$	0.940(0.060)	1.039(0.039)	0.004	0.002
		$\hat{\beta}_2$	-0.986(0.020)	-1.075(0.075)	0.001	0.006

The estimate mean values are close to true parameter values. Also, the mean absolute bias (MAB) and the MSEp for both the Cox model and DTLM are decreasing with an increase in sample size but the two events (Lognormal and Weibull) for the Cox model are underestimated with large MSEp.

For the Overall Event, we have precise estimates and there is no loss of efficiency in terms of MSEp. DTLM are over-estimated while Cox model are underestimated. Considering the two models for the estimates, Discrete-time model performs better than Cox model.

Table 3: Estimates, Absolute Bias and Mean Squared Errors for Weibull-Lognormal mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$)
Event 1 (Weibull), Event 2 (Lognormal)

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1 (Weibull)	100	$\hat{\beta}_1$	0.972(0.097)	1.027(0.085)	0.015	0.012
		$\hat{\beta}_2$	-0.935(0.163)	-0.961(0.167)	0.042	0.044
	200	$\hat{\beta}_1$	0.899(0.111)	0.915(0.099)	0.017	0.014
		$\hat{\beta}_2$	-0.935(0.163)	-0.961(0.167)	0.042	0.044
	500	$\hat{\beta}_1$	0.874(0.126)	0.981(0.045)	0.018	0.003
		$\hat{\beta}_2$	-0.944(0.087)	-1.048(0.092)	0.012	0.013
	1000	$\hat{\beta}_1$	0.854(0.146)	0.966(0.042)	0.023	0.003
		$\hat{\beta}_2$	-0.741(0.259)	-0.829(0.172)	0.071	0.034
	2000	$\hat{\beta}_1$	0.865(0.135)	0.997(0.022)	0.019	0.001
		$\hat{\beta}_2$	-0.864(0.136)	-0.972(0.044)	0.020	0.003
Event 2	100	$\hat{\beta}_1$	0.830(0.177)	0.898(0.123)	0.042	0.022
		$\hat{\beta}_2$	-0.849(0.197)	-0.898(0.183)	0.059	0.053
	200	$\hat{\beta}_1$	0.742(0.258)	0.807(0.194)	0.074	0.043
		$\hat{\beta}_2$	0.895(0.144)	-1.016(0.121)	0.031	0.023
	500	$\hat{\beta}_1$	0.747(0.253)	0.896(0.105)	0.067	0.014
		$\hat{\beta}_2$	-0.856(0.148)	-1.025(0.080)	0.029	0.010
	1000	$\hat{\beta}_1$	0.754(0.246)	0.895(0.105)	0.062	0.012
		$\hat{\beta}_2$	-0.650(0.350)	-0.761(0.239)	0.126	0.061
	2000	$\hat{\beta}_1$	0.773(0.227)	0.941(0.059)	0.052	0.004
		$\hat{\beta}_2$	-0.775(0.225)	-0.929(0.074)	0.053	0.007
Overall Event	100	$\hat{\beta}_1$	0.899(0.106)	0.994(0.057)	0.016	0.005
		$\hat{\beta}_2$	-0.881(0.136)	-0.951(0.107)	0.027	0.019
	200	$\hat{\beta}_1$	0.811(0.189)	0.922(0.081)	0.039	0.009
		$\hat{\beta}_2$	-0.989(0.070)	-1.183(0.186)	0.008	0.043
	500	$\hat{\beta}_1$	0.824(0.176)	1.031(0.041)	0.032	0.003
		$\hat{\beta}_2$	-0.932(0.073)	-1.175(0.176)	0.008	0.035
	1000	$\hat{\beta}_1$	0.821(0.179)	1.029(0.033)	0.033	0.002
		$\hat{\beta}_2$	-0.711(0.289)	-0.878(0.122)	0.085	0.017
	2000	$\hat{\beta}_1$	0.837(0.163)	1.077(0.077)	0.027	0.006
		$\hat{\beta}_2$	-0.842(0.158)	-1.070(0.070)	0.026	0.006

The results in Table 5 above reveal that estimate mean values are precisely close to the true values. Also, mean absolute bias (MAB) and the MSEp are decreasing as a result of increase in sample size except the estimates and mean absolute bias (MAB) for Event2 (Gompertz) under the discrete-time model that are increasing. The Cox model are underestimated with large MSEp.
Considering the Overall Event, the estimated mean values are close to the true values but DTLM overestimated. Comparing the estimates of the two models, Discrete-time model performs well than Cox model.

Table 4: Estimates, Absolute Bias and Mean Squared Errors for Exponential-Gompertz mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$)
Event 1 (Exponential), Event 2 (Gompertz)

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1	100	$\hat{\beta}_1$	0.965(0.086)	0.992(0.076)	0.011	0.009
		$\hat{\beta}_2$	-1.104(0.170)	-1.137(0.176)	0.046	0.048
	200	$\hat{\beta}_1$	1.003(0.055)	1.154(0.155)	0.005	0.029
		$\hat{\beta}_2$	-0.821(0.187)	-0.956(0.106)	0.047	0.018
	500	$\hat{\beta}_1$	1.096(0.097)	1.131(0.131)	0.012	0.019
		$\hat{\beta}_2$	-0.860(0.143)	-0.868(0.136)	0.026	0.024
	1000	$\hat{\beta}_1$	0.945(0.057)	1.031(0.037)	0.004	0.002
		$\hat{\beta}_2$	-1.019(0.047)	-1.107(0.108)	0.003	0.014
	2000	$\hat{\beta}_1$	0.915(0.085)	0.984(0.022)	0.008	0.001
		$\hat{\beta}_2$	-0.960(0.046)	-1.020(0.034)	0.003	0.002
Event 2	100	$\hat{\beta}_1$	0.957(0.091)	0.985(0.080)	0.013	0.010
		$\hat{\beta}_2$	-1.111(0.178)	-1.149(0.190)	0.050	0.053
	200	$\hat{\beta}_1$	1.007(0.055)	1.155(0.156)	0.005	0.029
		$\hat{\beta}_2$	-0.823(0.186)	-0.955(0.106)	0.046	0.018
	500	$\hat{\beta}_1$	1.093(0.095)	1.124(0.124)	0.011	0.018
		$\hat{\beta}_2$	-0.861(0.141)	-0.866(0.137)	0.026	0.024
	1000	$\hat{\beta}_1$	0.944(0.057)	1.031(0.037)	0.004	0.002
		$\hat{\beta}_2$	-1.023(0.047)	-1.111(0.112)	0.003	0.015
	2000	$\hat{\beta}_1$	0.914(0.086)	0.982(0.023)	0.008	0.001
		$\hat{\beta}_2$	-0.963(0.044)	-1.023(0.036)	0.003	0.002
Overall Event	100	$\hat{\beta}_1$	0.979(0.047)	1.026(0.049)	0.004	0.004
		$\hat{\beta}_2$	-1.129(0.140)	-1.192(0.194)	0.027	0.046
	200	$\hat{\beta}_1$	1.026(0.037)	1.217(0.217)	0.002	0.049
		$\hat{\beta}_2$	-0.839(0.162)	-1.005(0.055)	0.030	0.005
	500	$\hat{\beta}_1$	1.126(0.126)	1.184(0.184)	0.017	0.035
		$\hat{\beta}_2$	-0.886(0.114)	-0.911(0.090)	0.015	0.010
	1000	$\hat{\beta}_1$	0.971(0.030)	1.089(0.088)	0.001	0.008
		$\hat{\beta}_2$	-1.047(0.049)	-1.167(0.167)	0.003	0.029
	2000	$\hat{\beta}_1$	0.940(0.060)	1.032(0.032)	0.004	0.001
		$\hat{\beta}_2$	-0.987(0.019)	-1.070(0.070)	0.001	0.005

In order to compare the performances for the different combinations of distributions for each parameter estimates with 200 sample sizes. The results are summarized in tables below for the estimates of mean value, mean absolute bias (MAB) and mean square error for prediction (MSEp) for Cox model and discrete-time logit model (DTLM).

Table 5: Estimates, Absolute Bias and Mean Squared Errors for Lognormal-Exponential mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$) Event 1 (Lognormal), Event 2 (Exponential),

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1	100	$\hat{\beta}_1$	1.017(0.109)	1.054(0.108)	0.019	0.019
		$\hat{\beta}_2$	-0.882(0.197)	-0.914(0.206)	0.061	0.066
	200	$\hat{\beta}_1$	0.827(0.175)	0.931(0.097)	0.038	0.014
		$\hat{\beta}_2$	-0.879(0.162)	-0.979(0.143)	0.039	0.032
	500	$\hat{\beta}_1$	0.813(0.187)	1.005(0.049)	0.038	0.004
		$\hat{\beta}_2$	-0.829(0.176)	-1.008(0.095)	0.040	0.014
	1000	$\hat{\beta}_1$	0.766(0.234)	0.969(0.044)	0.056	0.003
		$\hat{\beta}_2$	-0.809(0.191)	-1.001(0.045)	0.041	0.007
	2000	$\hat{\beta}_1$	0.764(0.236)	0.987(0.028)	0.056	0.001
		$\hat{\beta}_2$	-0.821(0.179)	-1.017(0.047)	0.034	0.003
Event 2	100	$\hat{\beta}_1$	1.037(0.109)	1.081(0.138)	0.019	0.031
		$\hat{\beta}_2$	-0.870(0.207)	-1.130(0.241)	0.067	0.089
	200	$\hat{\beta}_1$	0.903(0.110)	0.971(0.087)	0.017	0.012
		$\hat{\beta}_2$	-0.974(0.129)	-1.070(0.168)	0.026	0.044
	500	$\hat{\beta}_1$	0.849(0.151)	1.070(0.077)	0.026	0.009
		$\hat{\beta}_2$	-0.816(0.186)	-0.921(0.113)	0.044	0.020
	1000	$\hat{\beta}_1$	0.883(0.217)	1.004(0.036)	0.048	0.002
		$\hat{\beta}_2$	-0.834(0.166)	-1.089(0.099)	0.032	0.014
	2000	$\hat{\beta}_1$	0.775(0.225)	1.043(0.045)	0.051	0.003
		$\hat{\beta}_2$	-0.836(0.164)	-1.027(0.050)	0.029	0.004
Overall Event	100	$\hat{\beta}_1$	1.074(0.098)	1.250(0.250)	0.014	0.072
		$\hat{\beta}_2$	-0.907(0.130)	-1.049(0.141)	0.027	0.030
	200	$\hat{\beta}_1$	0.892(0.110)	1.157(0.158)	0.016	0.030
		$\hat{\beta}_2$	-0.957(0.087)	-1.251(0.253)	0.012	0.008
	500	$\hat{\beta}_1$	0.866(0.134)	1.257(0.257)	0.020	0.069
		$\hat{\beta}_2$	-0.881(0.121)	-1.267(0.267)	0.018	0.079
	1000	$\hat{\beta}_1$	0.821(0.179)	1.227(0.227)	0.033	0.053
		$\hat{\beta}_2$	-0.878(0.122)	-1.289(0.289)	0.017	0.087
	2000	$\hat{\beta}_1$	0.811(0.189)	1.235(0.235)	0.036	0.056
		$\hat{\beta}_2$	-0.880(0.120)	-1.291(0.291)	0.015	0.086

Table 6: Estimates, Absolute Bias and Mean Squared Errors for Lognormal-Gompertz mixture of distributions as baseline hazards ($\beta_1 = 1$ and $\beta_2 = -1$)
Event 1 (Lognormal), Event 2 (Exponential)

Model Effect estimates			$\hat{\beta}$ (MAB)		MSE_p	
	Sample Size/Model		Cox	Discrete-Time	Cox	Discrete-Time
Event 1	100	$\hat{\beta}_1$	1.017(0.109)	1.052(0.107)	0.019	0.018
		$\hat{\beta}_2$	-0.882(0.197)	-0.912(0.206)	0.061	0.066
	200	$\hat{\beta}_1$	0.827(0.175)	0.931(0.097)	0.038	0.014
		$\hat{\beta}_2$	-0.879(0.162)	-0.979(0.143)	0.039	0.032
	500	$\hat{\beta}_1$	0.813(0.187)	1.004(0.049)	0.038	0.004
		$\hat{\beta}_2$	-0.829(0.176)	-1.008(0.095)	0.040	0.014
	1000	$\hat{\beta}_1$	0.766(0.234)	0.967(0.045)	0.056	0.003
		$\hat{\beta}_2$	-0.810(0.190)	-1.001(0.065)	0.041	0.006
	2000	$\hat{\beta}_1$	0.764(0.236)	0.986(0.028)	0.056	0.001
		$\hat{\beta}_2$	-0.822(0.178)	-1.017(0.046)	0.034	0.003
Event 2	100	$\hat{\beta}_1$	1.037(0.109)	1.187(0.197)	0.019	0.055
		$\hat{\beta}_2$	-0.870(0.207)	-0.968(0.209)	0.067	0.070
	200	$\hat{\beta}_1$	0.903(0.110)	1.018(0.079)	0.017	0.010
		$\hat{\beta}_2$	-0.974(0.129)	-1.091(0.162)	0.026	0.043
	500	$\hat{\beta}_1$	0.846(0.154)	1.063(0.074)	0.026	0.008
		$\hat{\beta}_2$	-0.813(0.190)	-0.988(0.094)	0.045	0.014
	1000	$\hat{\beta}_1$	0.783(0.217)	1.006(0.037)	0.048	0.002
		$\hat{\beta}_2$	-0.835(0.165)	-1.050(0.078)	0.032	0.010
	2000	$\hat{\beta}_1$	0.771(0.229)	1.018(0.029)	0.053	0.001
		$\hat{\beta}_2$	-0.838(0.162)	-1.058(0.068)	0.029	0.007
Overall Event	100	$\hat{\beta}_1$	1.074(0.098)	1.248(0.248)	0.014	0.072
		$\hat{\beta}_2$	-0.907(0.130)	-1.047(0.140)	0.027	0.030
	200	$\hat{\beta}_1$	0.892(0.110)	1.157(0.158)	0.016	0.030
		$\hat{\beta}_2$	-0.957(0.087)	-1.251(0.253)	0.012	0.008
	500	$\hat{\beta}_1$	0.866(0.134)	1.255(0.255)	0.020	0.068
		$\hat{\beta}_2$	-0.880(0.122)	-1.265(0.265)	0.019	0.077
	1000	$\hat{\beta}_1$	0.821(0.179)	1.225(0.225)	0.033	0.052
		$\hat{\beta}_2$	-0.879(0.121)	-1.289(0.289)	0.016	0.087
	2000	$\hat{\beta}_1$	0.810(0.190)	1.232(0.232)	0.036	0.055
		$\hat{\beta}_2$	-0.881(0.119)	-1.291(0.291)	0.015	0.086

Table 7: Estimates, Absolute Bias and Mean Squared Errors for the different combination of distributions with ($\beta_1 = 1, n = 200$)

Model	Estimate	Combination of Distributions					
		W-E	W-L	W-G	E-G	L-E	L-G
Cox	$\hat{\beta}(MAB)$	1.021(0.035)	0.811(0.189)	1.026(0.037)	1.026(0.037)	0.892(0.110)	0.892(0.110)
	MSEp	0.002	0.039	0.002	0.002	0.016	0.016
DTLM	$\hat{\beta}(MAB)$	1.226(0.226)	0.922(0.081)	1.225(0.225)	1.217(0.217)	1.157(0.158)	1.157(0.158)
	MSEp	0.053	0.009	0.052	0.049	0.030	0.030

Table 8: Estimates, Absolute Bias and Mean Squared Errors for the different combination of distributions with ($\beta_2 = -1, n = 200$)

Model	Estimate	Combination of Distributions					
		W-E	W-L	W-G	E-G	L-E	L-G
Cox	$\hat{\beta}(MAB)$	-0.833(0.168)	-0.989(0.070)	-0.836(0.164)	-0.839(0.162)	-0.957(0.087)	-0.957(0.087)
	MSEp	0.032	0.008	0.031	0.030	0.012	0.012
DTLM	$\hat{\beta}(MAB)$	-1.011(0.056)	-1.183(0.186)	-1.009(0.056)	-1.005(0.055)	-1.2551(0.253)	-1.251(0.253)
	MSEp	0.005	0.043	0.005	0.005	0.080	0.080

5. DISCUSSION AND CONCLUSION

Based on Tables 1-6, we compared the Cox model and Discrete-time Logit model(DTLM) with competing risk marginally and jointly under the mixture of different baseline hazard distribution functions and sample sizes. The results reveal that estimated mean values are close to the true parameter values. The mean of the estimated values of small sample sizes compared with those of large sample sizes indicate overestimation when the sample size is small which means the smaller the sample size, the greater impact of covariates. The performance improved mostly among small sample sizes for both Cox and Discrete-time logit models.

The mean absolute bias (MAB) and MSEp decreased as a result of increase in sample size. In terms of precision, it can be remarked in the estimates that Discrete-time logit model exhibit less mean absolute bias (MAB) although, the MSEp were slightly higher than the Cox model. The noticeable pattern lies in MSEp, larger sample sizes indicate lower MSEp.

The Overall Event (mixture of the baseline distributions) provides precise estimates in terms of mean estimates; mean absolute bias (MAB) and MSEp that are more convergent to the true value of the parameters than when each event follows individual baseline distribution.

In comparison between the mixture of the distributions for the Overall Event, WE (Weibull-Exponential), EG (Exponential-Gompertz) and WG (Weibull-Gompertz) gives a precise estimates with minimum MSEp. In respect to the baseline hazard distribution, all the mixture of baseline hazard with Lognormal distribution that has no property of proportional hazard gives over-estimated values.

Comparing the estimates of the two models, estimated mean values of covariate effects with the Cox model were obviously lower than the discrete-time logit model likewise in terms of mixture of baseline distributions, DTLM gives precise estimates.

References

Allison P. (1982). Discrete-time methods for the analysis of event histories sociology methodology. Doi: 10.2307/270718, 13:61-98.

Allison P. (2010) *Survival Analysis using SAS:A practical Guide*. volume 2.

Altman, D. G., De Stavola, B. L., Love, S. B., and Stepniewska, K. A., 1995. Review of survival analyses published in cancer journals. *Br J Cancer*. 72(2):511-518.

Anderson, G., and Fleming, T., 1995. Model misspecification in proportional hazards regression. biometrika 82: 527-541. *African Journal of Reproductive Health*. 82:527- 541.

Barber, J. S., Murphy, S. A., Axinn, W. G., and Maples, J., 2000. Discrete-time multilevel hazard analysis. *journal of Sociological Methodology*.30.

Bender, R., Thomas, A., and Blettner., M., 2005. Generating survival times to simulate cox proportional hazards models. *Statistics in Medicine*. 24:1713-1723.

Collect, D. (1994) Modelling survival data in medical research. *Chapman and Hall, London*.9:243-260.

Cox, D. R. (1972) Regression models and life-tables (with discussion). *Journal of Royal statistical society*. Series B. 34:187-220.

David, M. and Albert, N. (2014) *R Package survsim for the Simulation of Simple and Complex Survival Data*. volume 59(2). Journal of Statistical Software.

Enderlein, G., Cox, D. R., and Oakes, D., 1987. Analysis of survival data. *Biometrical Journal*. 29(1):114.

Hosmer, D. and Lemeshow, S., 1999. *Applied Survival Analysis:Regression Modelling of time to event data*.

Hougaard, P. (2000). *Analysis of multivariate Survival Data*. 560pp.

Kleinbaum, D. G. and Klein, M. (2012). *Survival analysis: A self-learning text*. volume 3.

McCallon, M. (2009). A discrete-time survival analysis of student departure from college. VDM Verlag.

Sharaf, T., and Tsokos, C. P. (2014). Predicting survival time of localized melanoma patients using discrete survival time method. *Journal of Modern Applied Statistical Methods*. 13(1).

Willett, J. B. and Singer, J. D. (1991). Modeling the days of our lives: using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin*Vol 110(2).

Willett, J. B. and Singer, J. D. (1993). Investigating onset, cessation, relapse, and recovery: why you should, and how you can, use discrete-time survival analysis to examine event occurrence.*J Consult Clin Psychol*. 61(6).

Singer, J. and Willett, J. (1993). Using discrete-time survival analysis to study duration and the timing of events. *Journal of Educational and Behavioral Statistics*.

Singer, J. D. and Willett, J. B. (1995). It's déjà vu all over again: Using multiple spell discrete-time survival analysis. *Journal of Educational and Behavioral Statistics*. 20(1).

Singer, J. and Willet, J. (2003). *Applied Longitudinal Data Analysis: Modelling Change and Event Occurence*.

Xie, H., McHugo, G., Drake, R., and Sengupta, A. (2003). Using discrete-time survival analysis to examine patterns of remission from substance use disorder among persons with severe mental illness. *journal of Ment Health Serv Res*. 5(1).