Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 10, Issue, 01, pp.63902-63906, January, 2018

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

MODELING TIME-TO-RECOVERY OF ADULT DIABETIC PATIENTS: A COMPARISON OF COX-PROPORTIONAL HAZARD AND SHARED GAMMA FRAILTY MODELS

***,[1]Abiyot Negash Terefe, [2]Dr. Yehenew Getachew and [3]Belay Birlie**

[1]Department of Statistics, College of Natural and Computational Science, Mizan-Tepi University, Tepi, Ethiopia
[2]Department of Plant Science, College of Veterinary medicine, Jimma University, Jimma, Ethiopia
[3]Department of Statistics, College of Natural and Computational Science, Jimma University, Jimma, Ethiopia

ARTICLE INFO

Article History:
Received 08th October, 2017
Received in revised form
20th November, 2017
Accepted 07th December, 2017
Published online 19th January, 2018

Key words:

Diabetic mellitus,
Shared gamma frailty model,
Cox-Snell residuals.

ABSTRACT

Background: The chances of developing diabetic may depend on a mix of genes, lifestyle and environmental factors. The objective of this study is modeling time-to-recovery of adult diabetic patients in a comparison of Cox-proportional hazard and shared gamma frailty models.
Methods: A retrospective data was obtained from Jimma University Specialized Hospital diabetic patient clinic. All diabetic patients 18 years of age and who are under treatments in between September 2010 and August 2013 are included in the study. Time of fasting blood sugar level to reach the first normal range, 70-130 mg/dl, of blood since time of treatment or intervention were the response variable. Due to the impact of residential places and unmeasured shared similarities in a cluster, district (Woreda) is used as a random effect (frailty) term in the survival models. In this study, Cox-PH and shared gamma frailty model were used.
Results: Types of diabetic, bodyweight at baseline, fasting blood sugar at baseline, sex and age of patients are significantly associated with time to recovery of diabetic patients. These variables are important factors that should be considered during the selection phase a treatment (combination of treatments) for diabetes.
Conclusion: The Cox-PH with shared gamma frailty model is the most powerful one in predicting recovery time of diabetic patients when there is significant difference among districts.

Citation: Abiyot Negash Terefe, Dr. Yehenew Getachew and Belay Birlie, 2018. "Modeling time-to-recovery of adult diabetic patients: a comparison of cox-proportional hazard and shared gamma frailty models", *International Journal of Current Research*, 10, (01), 63902-63906.

INTRODUCTION

The chances of developing diabetic may depend on a mix of genes, lifestyle and environmental factors. Environmental factors that contribute to beta cell destruction and genes regulating immune response are involved. Numerous environmental events trigger the autoimmune process in genetically susceptible individuals. There are environmental factors which have a link with DM like, chemical compounds (rodenticides, heavy metals virus, rarely and exposure to bovine milk proteins), and physical factors (penetrative short-wave length rage)etc. People with underlying medical conditions such as diabetes are more vulnerable to the adverse health impacts of climate change. In hotter temperatures, dehydration and heatstroke increases morbidity and mortality in people with diabetes. People with diabetes are predisposed to cardiovascular events during heat waves and higher mortality from heart attack on days of high air pollution (Dereje, 2005).

***Corresponding author: Abiyot Negash Terefe,**
Department of Statistics, College of Natural and Computational Science, Mizan-Tepi University, Tepi, Ethiopia.

The world health organization (WHO) publishes standards of medical care yearly to promote the importance of achieving optimal glycaemic control. Diabetes was classified according to WHO recommendations. Recommended blood sugar for people with diabetes (according to the WHO) before meals plasma glucose levels within a narrow range 70-130 mg/dl (milligram per deciliter). The statistical analysis of survival data is an important topic in many areas, including medicine, epidemiology, biology, demography, economics, engineering and other fields. A variety of techniques have been developed to analyse survival data. A common approach to the analysis of survival data is based on the assumption that the study population is homogeneous. That is, conditional on the covariates, every individual has the same risk of experiencing an event such as death or disease recurrence (Ulviya, 2013). The event times of individuals in the population, conditional on the observed covariates, are assumed to be independent. However, this cannot be assumed in all applications as many applications require heterogeneous sample, i.e. individuals with different risks and hazards. In practice, there may be an association between the events times of some subgroups of the population since the individuals of these groups share a

common trait that cannot be observed. For example, there may be an association in the times to events of cancer or cardiovascular diseases between siblings or married couples, even occurrence of nonlethal diseases within the same individual. Though individuals may look identical in some aspects, they may differ in unmeasured ways. In applications of survival analysis, usually only a few covariates such as age, sex, severity of disease or laboratory data are known. It is known that there are many other factors that can influence survival, including health status, life style, smoking, occupation and genetic risk factors. These factors are unknown and cannot be included in the analysis. A random effects model suggested in order to account for the unobserved heterogeneity due to unobserved covariates (Beard, 1959; Vaupel *et al.,* 1979; Lancaster *et al.,* 1979). The term longevity factor used to improve the effect of mortality models in populations (Beard, 1959). The term frailty introduced in order to account for unobserved heterogeneity, random effects, and association in univariate survival models (Vaupel *et al.,* 1979). He introduced this concept of frailty to biostatistics by applying it on population mortality data. The model introduced to the literature of economics and the model is called the mixed proportional hazards model (Lancaster *et al.,* 1979). The concept, however, goes back to work of (Green, 1987) on "accident proneness" in 1920. The applications of the model to multivariable survival data discussed in his seminar paper on chronic disease incidence in families (Clayton and Cuzick, 1985). Frailty models account for unobserved heterogeneity that occurs because some observations are more prone to failure, and therefore more "frail" than others in a data set. Therefore, the objective is to introduce an additional parameter to the hazard rate that accounts for the random frailties. These frailties can be specific to groups, and are referred to as shared frailty. The overall aim of this thesis is modeling of time-to-first recovery of adult diabetic patients from Jimma University Specialized Hospital (JUSH) using various survival models.

Inference for Cox PH model (Cox, 1972) was developed under the assumption that the observations are statistically independent and the population they come from is assumed to be homogeneous with respect to failure. However, this assumption may be violated. Thus, in many epidemiological studies, failure times are clustered into groups such as families or geographical units; some unmeasured characteristics shared by the members of that cluster, such as genetic information or common environmental exposures could influence time to the studied event. In a different context, correlated data may come from recurrent events, i.e. events which occur several times within the same subject during the period of observation. Ignoring the existence of heterogeneity will produce incorrect estimation of parameters and their standard errors in survival analysis. Ignoring heterogeneity overestimates life expectancy based on their study on estimating life expectancy in a heterogeneous population (Keyfitz and Littman, 1979). Ignoring frailty leads to regression coefficient estimates biased towards zero by an amount depending on the distribution and the variability of the frailty terms (Henderson and Oman, 1999). For such situations, one approach accounting for correlation is to incorporate an additive or multiplicative random effect for each cluster, resulting in a frailty model. Random effect or frailty model attempts to account for the existence of unmeasured attributes (such as genotype, environment and geographical location) that introduce heterogeneity into the study population. Not taking into account the unobserved frailty will thus under/overestimate the model parameters. Therefore, in this thesis, we are interested to address the following interesting research questions; Which type of DM takes long time to recover to normal blood sugar level; is there heterogeneity among districts with respect to time to recovery of adult diabetic patients; what are the covariates influencing the time to recovery for each type of diabetic; and the general objective of the study is modeling time-to-first recovery of adult diabetic patients in a comparison of Cox-Proportional Hazard and Shared Gamma Frailty Models shared gamma frailty Model. The results of this study will be very useful in the development of an effective diabetic care and anti-diabetic therapy (ADT) patient monitoring system.

METHODS

Longitudinal retrospective cohort follow up of adult diabetic patients is collected from JUSH Diabetic Patient Clinic located in southwest of Ethiopia. The data is extracted from the patient's chart which contains epidemiological, laboratory and clinical information of all diabetic patients under insulin treatment follow-up. A total of 1930 diabetic patients are on active follow up. All diabetic patients greater than or equal to 18 years old and placed under treatments that have followed between September 2010 and August 2013 (three years data) were included. The data for this study consists of 544 individuals. The outcome variable considered in this study is the time to first recovery of diabetic patients until it reaches normal fasting (before meal) blood sugar level in the follow up period. Covariates considered are Sex, Age (years), Diabetic type, Family history, systolic and diastolic blood pressure in mm/hg, Bodyweight in kg, and fasting blood sugar in mg/dl. Model building is the process of developing a probabilistic model that best describes the relationship between the dependent and independent variables.

Survival Analysis

The Cox proportional hazards (PH) regression model (Cox, 1972) a broadly applicable and the most widely used method of survival analysis.

$$h_i(t) = h_0(t)\mu_t = h_0(t)\exp(X_i^T\beta) \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where, $h_0(t)$, is the baseline hazard function; X_i is a vector of covariates and β is a vector of parameters for fixed effects. Parameter estimate β refers to the increase in log-hazard with a one unit increase for the continuous covariate.

Frailty Model

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and unobserved heterogeneity (Vaupel *et al.,* 1979). First frailty term is introduced, promoted the model by its application to multivariate situation on chronic disease incidence in families (Clayton and Cuzick, 1985). The random effect, called frailty and denoted here by *Z* is the term that describes the common risk, acting as a factor on the hazard function.

Shared Gamma Frailty model

Conditional on frailties, Z_i, the survival times are assumed to be independent and their hazard functions to be of the form

$$h(t_{ij}/Z_i) = Z_i h_0(t_{ij})\exp(\beta^T X_{ij}), i = 1,2,\ldots n; j = 1,2,\ldots,k_i \quad ..(2)$$

Where, $h_0(t)$ are the baseline hazard functions and β is a vector of fixed effect parameters to be estimated. The frailties Z_i are assumed to be identically and independently distributed random variables with a common density function $f(z,\theta)$ where, θ is the parameter of the frailty distribution. It is assumed that the Z_i's are independently and identically distributed from a gamma distribution with mean 1 and unknown variance θ; the probability density function is thus:

$$f_z(z) = \frac{z^{\frac{1}{\theta}-1}\exp(-z/\theta)}{\theta^{\frac{1}{\theta}}\Gamma(\frac{1}{\theta})} \quad \ldots\ldots\ldots\ldots\ldots..(3)$$

Where, $\Gamma(.)$ is gamma function, it corresponds to a gamma distribution, $Gamma(\mu,\theta)$ with $\Gamma(.)$ μ fixed to 1 for identifiability. Its variance is then θ with Laplace transform

$$L(u) = (1+u/\theta)^{-\theta},\ u \geq 0$$

In this model, the estimates of coefficients are obtained by using the penalized likelihood maximization (PLM). Parameter estimates were obtained by Breslow approximation to partial likelihood. After fitting the models to a set of survival data, the adequacy of the fitted models to the survival data were checked using Cox-Snell residuals (Hosmer and Lemeshow, 1999; Collet, 1972).

RESULTS

Descriptive Results of patients following anti-DM treatment

From Table 1 out of 544 DM patients, 404 (74.26%) were recovered and the rest 140 (25.74%) censored from the study. The median recovery time of the patients was 3 months but it varies depending on the covariates included in the study. Amongst the patients infected with DM, 193 (35.48%) were female and 351 (64.52%) were male. Regarding the patients infected with DM, 121(22.24%) were type-I DM and 423(77.76%) were type-II DM. Among patients, 423(77.76%) were not from family history and 121(22.24%) were from family history. The median recovery time for female and male DM patients were 6 and 2months respectively. As compared to male patients majorities of female patients recovered the disease.

Multivariable Cox-PH and Shared Gamma Frailty Models

In order to select variables in the model, first univariable analysis is used to check all the covariates associated with recovery time. The multivariate results of a Cox PH model fitted to this dataset were obtained on Table 3. It is now observed that effects of age group, bodyweight (kg) (p-value= 0.000) at baseline, sex (p-value=0.000), FBS (mg/dl) (p-value=0.000) at baseline and diabetic type (p-value= 0.000) are significantly associated. The multivariate results of a Cox PH with gamma frailty model (2) fitted to this dataset was obtained on table 3. In this table, all covariates are statistically significant at 5% level of significance. In Cox-PH model the estimate for β and its estimated error is smaller compared to the shared gamma frailty models (2). The LR of a Cox-PH model without frailty (*log likelihood*=-2061.395) and with gamma frailty model in frailty (*I-likelihood*=-2053) is 2(2061.395-2053) =16.79, its' p-value=2.088e-05, there is a significant frailty effect, implies correlation within district cannot be ignored. In gamma frailty models indicating that frailty variable (districts) is very highly significantly related to the time to first recovery of DM. Thus, there is much evidence pointing towards a population that is indicating heterogeneity. This might be due to the shared environmental and residential factors. Hence, Cox PH model with shared gamma frailty model provide a suitable choice for modeling time to first recovery of DM as compared to Cox PH model.

Table 1. Descriptive Baseline of DM patients for categorical variables

Covariates		Total No (%)	Recovery	Median	95% CI
Diabetic Type	Type-I	121	103	2	(1, 2)
	Type-II	423	301	4	(4, 5)
Sex	Male	351	273	2	(2, 3)
	Female	193	131	6	(5,8)
Family history	Yes	121	103	2	(1, 2)
	No	423	301	4	(4, 5)
Over all DM		544	404	3	(3,4)

Table 2. Cluster Variance Estimate

Between-Cluster Variance Estimate		
Cluster	Theta	Standard error
District	0.18	0.081

After controlling for other prognostic factors and accounting for frailty in table 3, patients with age group 30-44, 45-59, 60-74 or >74 are recovered 0.679, 0.525, 0.273 and 0.159 times age group of 18-29. Being young is associated with better recovery. The results reveal that after accounting for heterogeneity and other confounders in the data, time to recovery takes longer time with a unit increase of FBS and bodyweight at baseline in diabetic patients. This implies that the lower the age, FBS and bodyweight at baseline the faster the rate of recovery (blood sugar level reaching the normal range) of diabetic patients. Type of diabetes has a significant effect on the life of diabetic patients. A hazard ratio of 0.582 indicates, 36.79% chance of the type-II diabetic patients recovered first as compared to type-I diabetic patients after accounting and controlling other factors in the model. Thus, an individual suffering from type-II diabetes delayed to recover as compared to type-I diabetic and also, sex is seen to be significantly associated with the recovery time of the diabetic patient. After accounting for heterogeneity and other factors, the male diabetic recovers 1.728times the female. That is, males recover faster than females (HR=1.728, 63.34% chance of the male diabetic patients recovered first).

The final model then given by:

$$h(t_{ij}/Z_i) = h_0(t_{ij})Z_i\exp(-0.387*Age_{30-44ij} - 0.645*Age_{45-59ij} - 1.299*Age_{60-74ij} - 1.839*Age_{>74ij} - 0.022*bodyweight_{ij} - 0.541*DT_{ij} - 0.0033*FBS_{ij} + 0.547*Sex_{ij})$$

and $$f_z(z) = \frac{z^{\frac{1}{0.0526}-1}\exp(-z/0.0526)}{0.0526^{\frac{1}{0.0526}}\Gamma(\frac{1}{0.0526})}$$

Table 3. Multivariable Cox-PH and shared gamma frailty models for the diabetic patients

Non- Frailty Model				Frailty Model		
Covariates	Hazard Ratio(HR)	se(Coef$_{(\hat{\beta})}$)	P-value	Hazard Ratio(HR)	se(Coef$_{(\hat{\beta})}$)	P-value
Age						
30-44 years	0.679	0.135	0.004	0.679	0.137	0.005
45-59 years	0.499	0.139	0.000	0.525	0.142	0.000
60-74 years	0.252	0.186	0.000	0.273	0.189	0.000
>74 years	0.148	0.347	0.000	0.159	0.351	0.000
Bodyweight(kg)	0.979	0.003	0.000	0.979	0.003	0.000
Diabetic Type (Type-II)	0.617	0.124	0.000	0.582	0.125	0.000
FBS	0.997	0.001	0.000	0.997	0.000	0.000
Sex (Male)	1.79	0.113	0.000	1.728	0.114	0.000
Likelihood ratio test= 251.2.1, p=0.000, Wald test = 222.2 p=0.000, Score (logrank) test = 243.3, p=0.000, AIC= 4138.79.				Theta (θ) =0.053, $\bar{\chi}_{01}^2$ = 16.79, p-value= 0.000, I-likelihood = -2053, Likelihood ratio test= 287, p=0.000, AIC= 4117.636		

*Inseparable: when two types of diabetic is in one data set, Coef: coefficient for covariate, HR: hazard ratio, p-value: probability value, 95%C.I HR: 95% confidence interval for HR, FBS: Fasting Blood Sugar, * Significant at 0.05 level.*

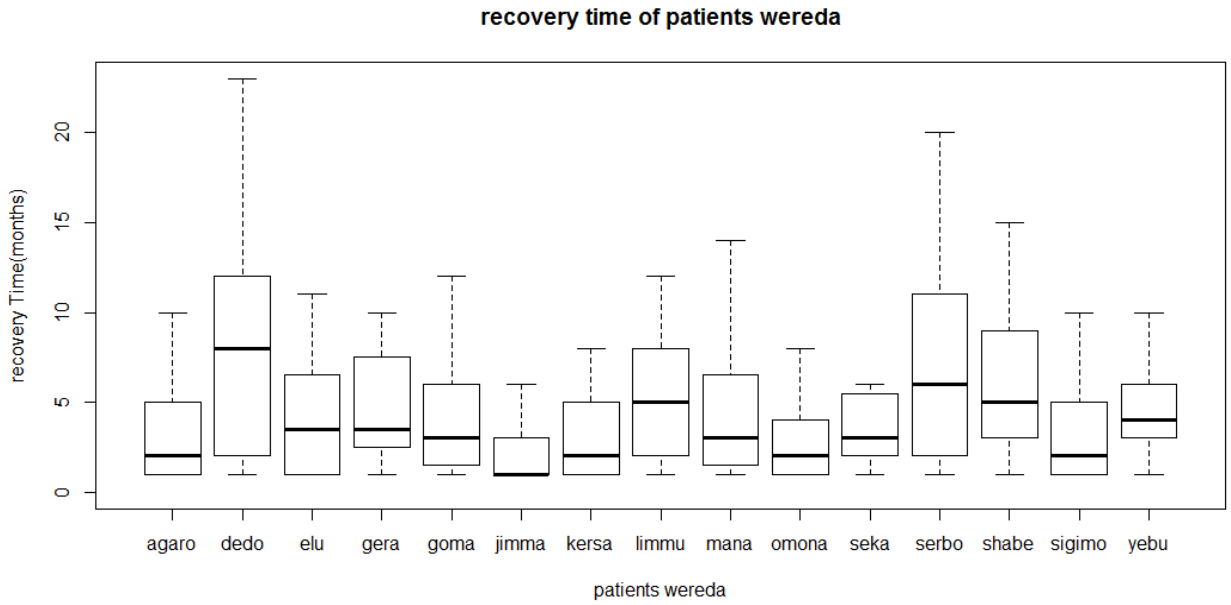

Figure 1. Box plot for recovery time (months) of patients District

Data reveals moderate dependence $(\theta = 0.0526, kendall's\ \tau = 0.02563)$. Since, Kendall's τ is 0.02563 for Cox-PH with shared gamma frailty, thus there is a positive correlation of 0.0263 between the recovery times of diabetic patients within district. From figure 1 below, the median first recovery time for different districts are significantly different, describe us presence of heterogeneity. Since, patients in dedo and serbo districts takes long time to recover, therefore, the median recovery times are delayed where as patients in Jimma district, they are fast to recover into normal blood sugar level.

Between-Cluster Variance Estimate

The results in table 2 show the estimated shared gamma frailty model with random effects. A large value of cluster variance (theta=0.18) indicates a greater degree of heterogeneity among districts and strong association within districts. From figure 2 the patients' chance of recovery time up to 15 months in districts slowly decreases and increases at baseline survival and hazard function respectively. A plot of figure 3 is the Cox-Snell residuals against the cumulative hazard function, follows the 45 degree line very closely except for very large values of time.

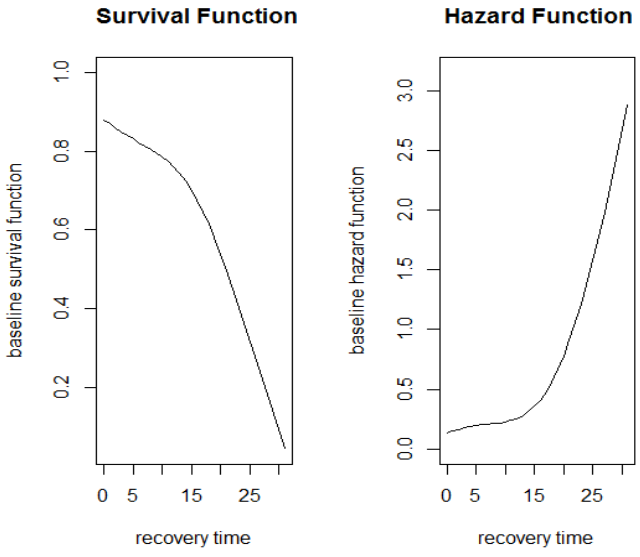

Figure 2. Estimated baseline survival and hazard in the PHs models of DM patients using Cox PH with shared gamma frailty model

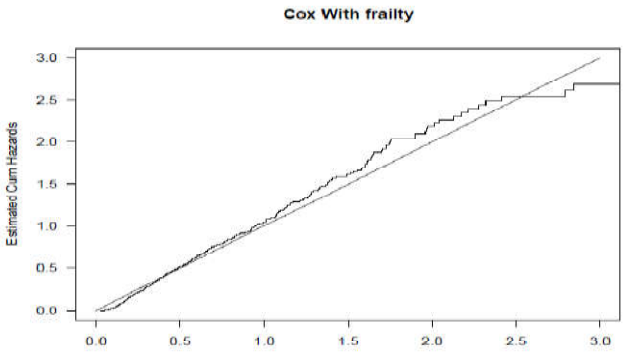

Figure 3. Cox-Snell residuals obtained from fitting Cox PH frailty model to the DM data

DISCUSSION

From the total of 544, 404(74.26%) experienced the event and the rest 140 (25.74%) loss to follow-up from the study. In this study first recovery time for different districts are significantly different, describe us presence of heterogeneity consistent with (Dereje, 2005) done in Ethiopian health center as well (Beard, 1959; Vaupel *et al.,* 1979; Lancaster, 1979) suggested a random effects model in order to account for the unobserved heterogeneity due to unobserved covariates. In Cox-PH with shared gamma frailty model; age group, bodyweight, diabetic type, FBS, and sex of patients at baseline shows a statistically significant association with time to first recovery to normal blood sugar level. In Cox-PH with frailty model, the types of diabetic was a strong and independent prognostic factor, indicating better recovery time for type-I patients controlling other factors in the model. This means that patients with type-II getting affected by diabetic mellitus prolonged recovery time as compared to type-I; these findings are consistent with those done in Uganda countries by (Olive *et al.,* 2007).

In Cox-PH with shared gamma frailty model (2) the recovery time of an individual suffering from type-II diabetes who has not recovered yet has 0.586 (HR=0.586, 36.95% chance of the type-II diabetic patients recovered first) times as compared to type-I DM. Being female prolonged the recovery time as compared to males (HR=2.026, 66.953% chance of the male diabetic patients recovered first) in type-II DM. In Cox-PH with shared gamma frailty model gender was a strong and independent prognostic factor, indicating males are better recovering to normal blood sugar as compared to females'. This means that females getting affected by diabetic mellitus (DM) have a slightly takes longer time to recover to normal blood glucose level than males, these findings are consistent with those again obtained in Ugandan countries. When the frailty is ignored, the estimate for β and its estimated error is smaller compared to the shared gamma frailty model (2). This is expected as the frailty model account for the extra variance associated with unmeasured risk factors (Keyfitz and Littman, 1979; Lancaster, 1979) showed that when heterogeneity is ignored, it caused underestimation of covariate effects in his study of unemployment rates and (Henderson and Oman, 1999), showed that ignoring frailty leads to regression coefficient estimates biased towards zero by an amount depending on the distribution and the variability of the frailty terms.

Conclusion

The median first recovery time for different districts are significantly different. Being old, female, higher FBS and overweight at baseline prolonged the recovery time; hence, concerned body should give special treatment for older, females, Higher FBS, Overweight, and type-II DM of patients at baseline. In Cox-PH model the estimate for β and its estimated error is smaller compared to the shared gamma frailty models (2). This is expected as the frailty model account for the extra variance associated with unmeasured risk factors. The Cox-PH with shared gamma frailty model is the most powerful one in predicting recovery time of diabetic patients when there is significant difference among districts.

Acknowledgment: We gratefully thanks to our parents, whose love, trust, encouragement, comfort and support always surround. Also, we really appreciate the precious learning opportunity, budget and environment provided by Jimma University. We gratefully thank to all the faculty, department, and friends who ever helped. Lastly not from least, I would like to thank Jimma University Specialized Hospital (JUSH) Diabetic Patient Clinic for their great help during data collection.

Conflict of interest statement: I have no conflict of interest to declare.

REFERENCES

Beard, R.E. 1959. Note on some mathematical mortality models, in: C.E.W. Wolstenholme, Connor (eds.. The life span of animals. CIBA Foundation Colloquia on Ageing.

Clayton, D., Cuzick, J. 1985. The semi-parametric Pareto model for regression analysis of survival times. Proceedings of the Centenary Session of the International Statistical Institute, Amsterdam.

Collet, D. 1972 Modeling Survival Data in medical research second edition.

Cox, 1972. Regression models and life tables (with discussion) *Journal of the Royal Statistical Society B*.

Dereje, A. 2005. Diabetes Mellitus, Diploma program for the Ethiopian health center team.

Green, P. J. 1987. Penalized likelihood for general semi-parametric regression models. International Statistical Review.

Henderson, R. and Oman, J. P. 1999. Effect of frailty on marginal regression estimates in survival. *Journal of the Royal Statistical Society*.

Hosmer, D.W. and Lemeshow, 1999. Applied survival analysis regression modeling of time to event data.

Keyfitz, N. and Littman, G. 1979. Mortality in a heterogeneous population.

Lancaster, T. 1979. Econometric Methods for the Duration of Unemployment. Econometrica.

Olive, D., Odwee, O. and Atuhaire, K. 2007. The determinants of recovery time of diabetic patients from three hospitals in Uganda, 2007.

Ulviya, A. 2013. Frailty Models for Modeling Heterogeneity.

Vaupel, J. W., Manton, K. G., and Stallard, E. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography.

****