Research on Humanities and Social Sciences
ISSN (Paper)2224-5766 ISSN (Online)2225-0484 (Online)
Vol.4, No.21, 2014
www.iiste.org

Performance of Cox Proportional Hazard and Accelerated Failure Time Models in the Analysis of HIV/TB Co-infection Survival Data

Derek Ngbandor Nawumbeni *, Albert Luguterah, Timothy Adampah
Department of Statistics, University for Development Studies, P. O. Box 24, Navrongo, Ghana
* E-mail of the corresponding author: nawumbenid@gmail.com

Abstract

Cox model and accelerated failure time models are widely used in the modeling of survival data for various diseases. This paper compares the performance of these two models viz. Cox Proportional Hazard Model and the Accelerated Failure Time Model using HIV/TB Co-infection Survival data. The study revealed that, the AFT model has the best predictive power compared to the Cox model based on the AIC and BIC values.

Keywords: Cox proportional hazard model, Accelerated failure time model, Cox-Snell residual, HIV/TB Co-infection, CD4 cell count

1. Introduction

Survival analysis is a method for analysing the occurrence of a given event in which individuals are followed from the time they experience the particular event such as the diagnosis of disease, and the time to recurrence of the disease or death (Kirkwood and Sterne, 2006). This event can be the development of a disease, treatment outcome, relapse, or death. Survival data can include survival time, outcome of treatment of a disease, and patient characteristics related to response, survival, and the development of a disease (Lee and Wang, 2003). Due to the limitations in survival data, the usual statistical methods cannot be used in survival analysis. The limitations include censoring, skewing and lack of normality in the distribution. The methods for analysing survival data include; nonparametric, semi-parametric and parametric. Some researchers prefer the Cox proportional hazard model as the appropriate model to analyse the survival data. However, the Accelerated Failure Time model can be more suitable in some instances.

Sayehmir *et al.,* (2008) studied the prognostic factors of survival time after hematopoietic stem cell transplant in acute lymphoblastic leukemia patients in Shariati Hospital, Tehran. Their study was between 1993 to 2007 using the Cox PH and accelerated failure time models. They concluded that the predictive power of Weibull AFT model was superior to Cox PH model.

Ponnuraja and Venkatesan (2010) in their study, the proportional hazard model and its extension were used comprehensively to assess the effect of an intervention in the presence of covariates. They observed that in situation where the effect of the intervention is to accelerate the PH assumptions may not hold hence the AFT model is also appropriate. Their study was aimed to formulate a model that yields biological plausible and interpretable estimates of the effect of important covariates on survival time. It was revealed that the AFT model gives better prediction than the Cox PH model.

Ravangard *et al.,* (2011) compared the Cox proportional hazard model and parametric models in studying the length of stay in a Tertiary Teaching Hospital in Tehran. The AIC and Cox-Snell residual graph showed that the Gamma (AFT) model fitted the data best.

Vallinayagam *et al.,* (2014) compared the performance of the common parametric models including the, Exponential, Weibull, Gompertz, Lognormal and Log-logistic using Breast Cancer data. Their study revealed that Log-normal model is better than other models. This research is very essential because it compares the Cox regression model and Accelerated Failure Time model in HIV/TB Co-infection which is limited in the available literature.

2. Materials and Methods

The study considers a real-life data set obtained from St. Mathias Hospital in the Pru District of the Brong Ahafo Region of Ghana. This hospital serves as a referral center for different health centers in the District. The hospital has a unit for both ART and TB. The hospital started giving free ART services in 2008. Data was extracted from

the patient folders, which have been adopted by the Ministry of Health, Ghana. The study considered patients with ages above five years. The study period was between the year 2008 to 2013 and the patients followed till the outcomes of either the event (failure) or censored.

2.1 Estimation of the survivorship function: We used the Life table method to estimate the survivorship function. The Gehan's method (1969) was employed where the midpoint of the interval was used to estimate the Hazard and the density functions and the upper limit used to estimate the survival function.

2.2 Log rank test: This was used to compare the death rate between two distinct groups, conditional on the number at risk in the groups. The log rank test hypothesis that;

H_0 : All survival curves are the same

H_1 : Not all survival curves are the same.

Log rank test approximates a chi-square test which compares the observed number of failures to the expected number of failure under the hypothesis.

$$\chi^2 = \sum_{f=1}^{k} \frac{(O_f - E_f)^2}{E_f} \qquad (1)$$

where, $k-1$ is the degrees of freedom. A large chi-squared value implies a rejection of the null hypothesis for the alternative hypothesis.

***2.3 Cox Proportional Hazard Model*:** The Cox proportional Hazard regression proposed by Cox (1972) is used to determine the multiplicative hazard of some prognostic factors on HIV/TB Co infection. The variable *X* represents a collection of predictor variables that is being modelled to predict the individual hazard of a patient. The Cox model is represented as:

$$h(t,X) = h_0(t) e^{\sum_{i=1}^{P} B_i X_i} \qquad (2)$$

where: $h_0(t)$ is the baseline hazard function, *e* is the exponential expression to the linear sum (this sum is over *p* explanatory variables), X_i is the explanatory or the predictor variable and β_i is the regression coefficient.

***2.4 Accelerated Failure Time Models (AFTM)*:** This model is assumed to follow a known distribution. The models include the: Exponential model, Weibull, Lognormal, Log-logistics and Gamma models. The underlying assumption for this model is that the effect of the covariate is multiplicative with respect to the survival time. The model regresses the natural logarithm of the survival time (*log t*) over the covariates. It is expressed as a linear function of the covariateS.

$$\log t = X_j \beta + z_j \qquad (3)$$

Where X_j is the vector of covariate, β is the vector of regression coefficient, z_j is the error term.

2.5 Model Diagnostics: The Cox model was checked to determine whether the model satisfies the proportionality assumption. The martingale residual plot was conducted among the continuous covariates to ascertain the linearity between the covariates and the survival time. The Cox-Snell residual plot was done to determine whether the AFT model is well fitted.

3. Results and Discussions

There were 76 patients on treatment of HIV/TB Co-infection from the year 2008 to 2013. The study indicated that the percentage of deaths among the patients was 32.9%. This agrees with the Interagency Coalition on AIDS and Development findings in 2010 that, up to 33% of all AIDS deaths worldwide can be attributed to TB. This could also be due to the difficulty in diagnosing the HIV patients of TB since HIV patients are more susceptible to contracting TB outside the lungs. The life table estimate indicates that about 7% of the patients [HR=0.066667] failed in the first month of the treatment as shown in **Table 1** and **Figure 1**. The 63rd month is the riskiest month [HR=0.333333] as approximately 33% of the patients failed.

In determining whether there is significant difference among different groups of the covariates, we employed the log rank test of equality as shown in **Table 2**. The test indicates that Religion shows a significant difference of

survival among the patients. However, covariates including Sex, Marital status, Disclosure to sexual partner and Drug regimen are insignificant.

The Cox PH model for the patients confirm that the weight and CD4 cell count are significant at 5% significance levels as shown in **Table 3**. Their estimated hazard ratios and *p*-value as [HR = 0.913, *p*-value =0.0071] and [HR =0.993, *p*-value =0.0020] respectively. This implies that an increase in the CD4 count of a patient will decrease the estimated hazard by 0.993 assuming that all covariates are constant. Also, a unit increase in the Weight of a patient will lower the risk of the patient by 0.913 assuming that all other covariates are held constant. The Single patient is also significant at 10% significance level [HR =0.029, *p*-value =0.0826]. This implies that a Single patient have his/her estimated hazard decreased by 97% compared to the widowed patient holding other factors constant.

From **Table 4,** there is enough evidence to conclude that the proportionality assumption is not violated since the *p*-values are statistically insignificant. This suggests that, the Cox proportional hazard model is appropriate. Thus, the covariate does not correlate with the survival time. The martingale residual plot for the three continuous covariates in the Cox model: Age, Weight and CD4 count in **Figure 2** revealed that the plots are linear and showed a correct functional form. The result does not show any trend and the resulting smoothed plots (LOESS) can be described as horizontal straight lines.

The Gamma model in **Table 5** revealed that weight, CD4 cell count and the Religious status of the patients are significant at 5%. However, Gender, Age, Marital status, Drug regimen, WHO Clinical Stage and Disclosure were statistically insignificant. Thus, a unit increase in the weight of a patient corresponds to an increase in the survival time since the time ratio is greater than 1, [$TR = e^{0.0809}$ =1.084]. Similarly, a unit increase in CD4 count of a patient would improve the estimated survival time [$TR = e^{0.0043}$ =1.004]. This agrees with Rafera (2012) where he asserts that the rate of dying among patients with higher weight and CD4 cell count in Ethiopia is proportionally lower compared to patients with lower Weight and CD4 count. Patients who practice Christianity and Islam will have a better survival [$TR = e^{0.0217}$=1.022] and [$TR = e^{0.0103}$ =1.010].

4. Conclusion

In this study, the Cox model and the Accelerated Failure Time model have been compared using HIV/TB Co-infection data. The result showed that there were 76 patients on treatment. The Cox model was fitted and diagnosed with the proportionality assumption satisfied. The martingale residual indicated that the model was linear. Comparing the Cox model with the AFT model based on the AIC and BIC showed that the Gamma model had the lowest value. The percentage of death among the patients was 32.9. It was also observed that weight, CD4 cell count and the Religion were significant determinants of the patient's survival at 5% significance level. The result revealed that the gamma model provided a better fit to the studied data than the Cox proportional hazards model. Hence, it is better for researchers of HIV/TB Co-infection to consider AFT model even if the proportionality assumption of the Cox model is satisfied.

5. Acknowledgement

The researchers are grateful to Ruth Polishuk for financing this paper and St. Mathias Hospital administrator and staff for their contributions for us to obtain the data.

References

Gehan, E. A., (1969). Estimating Survival Function from Life Table. *Journal of chronic diseases,* 21: 629-644.

Interagency coalition on AIDS and Development, (2010). Available at www.icad-cisd.com

Kirkwood, B. R., and Sterne, J.A.C., (2006). Essential Medical Statistics, 2nd ed. Library of Congress Cataloging-in-Publication Data, British Library.

Lee, E. T., and Wang, J. W. (2003): Statistical Methods for Survival Data Analysis: John Wiley and Sons. New Jersey.

Ponnuraja, C., and Venkatesan, P (2010). Survival Models for Exploring Tuberculosis Clinical Trial Data-an Empirical Comparison. *Indian Journal of Science and Technology,* 3(7): 0974- 6846.

Ravangard, R., Arab M., Rashidian, A.,Akbarisari, A., Zare A., Zeraat, H. (2011). Compared Cox Model and

Parametric Models in the Study of Length of Stay in a Tertiary Teaching Hospital in Tehran, Iran. *Acta MedicaIranica*, 49(10): 650-658.

Refera, H., (2012). Survival and risk factors of HIV/TB Co-infected Patients under Antiretroviral Therapy in Ambo Hospital, Ethiopia. Addis Ababa University, School of Graduate Studies Department of Statistics.

Sayehmiri, K., Eshraghian R. M., Mohammad K., Alimoghaddam K., Foroushani R. A., Zeraati H., Golestan, A., Ghavamzadeh, A. (2008). Prognostic Factors of Survival time after Hematopoietic Stem Cell Transplant in acute Lymphoblastic Leukemia Patients in Shariati Hospital, Tehran. *Journal of Experimental & Clinical Cancer Research,* **27**(1): 1-9.

Vallinayagam, V., Prathap, S., and Venkatesan, P. (2014). Parametric Regression Models in the Analysis of Breast Cancer Survival Data. *International Journal of Science and Technology,* 3(3) 2049-7318.

APPENDIX

Tables and Figures of HIV/TB Co-infection

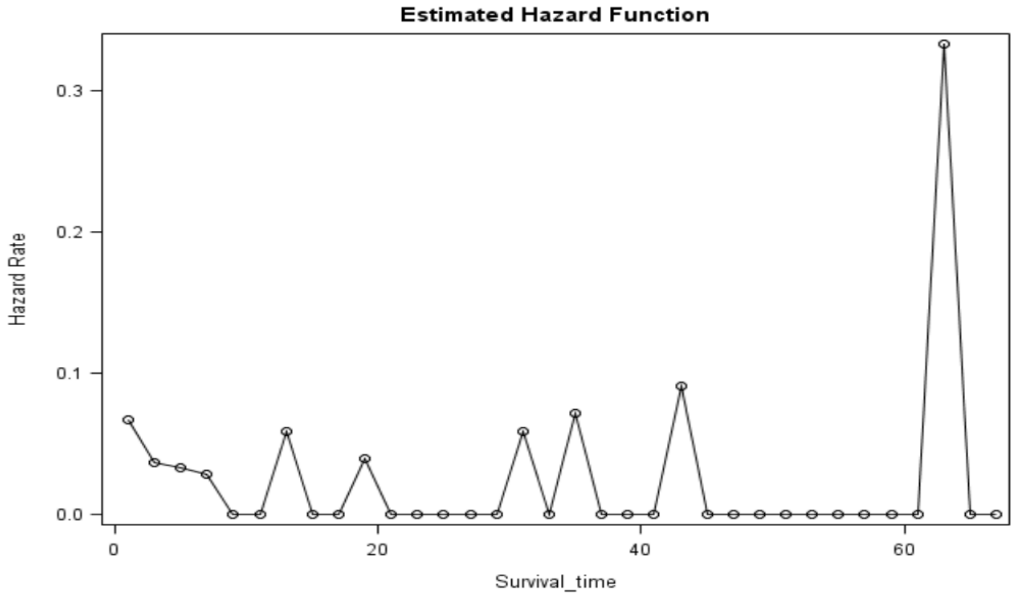

Figure 1: Hazard curve for co-infection

Table 1: Hazard, density and survival estimates for co-infected patients on treatment

Mid-point	Hazard	SE	Density	SE	Upper-limit	Survival	SE
1	0.0667	0.0222	0.0625	0.0195	2	1.0000	0.0000
3	0.0367	0.0183	0.0310	0.0150	4	0.8750	0.0390
5	0.0333	0.0192	0.0262	0.0147	6	0.8131	0.0469
7	0.0286	0.0202	0.0211	0.0146	8	0.7606	0.0528
9	0.0000	0.0000	0.0000	0.0000	10	0.7183	0.0577
11	0.0000	0.0000	0.0000	0.0000	12	0.7183	0.0577
13	0.0589	0.0415	0.0399	0.0268	14	0.7183	0.0577
15	0.0000	0.0000	0.0000	0.0000	16	0.6385	0.0739
17	0.0000	0.0000	0.0000	0.0000	18	0.6385	0.0739
19	0.0400	0.0340	0.0246	0.0238	20	0.6385	0.0739
21	0.0000	0.0000	0.0000	0.0000	22	0.5894	0.0829

23	0.0000	0.0000	0.0000	0.0000	24	0.5894	0.0829
25	0.0000	0.0000	0.0000	0.0000	26	0.5894	0.0829
27	0.0000	0.0000	0.0000	0.0000	28	0.5894	0.0829
29	0.0000	0.0000	0.0000	0.0000	30	0.5894	0.0829
31	0.0588	0.0587	0.0327	0.0312	32	0.5894	0.0829
33	0.0000	0.0000	0.0000	0.0000	34	0.5239	0.0962
35	0.0714	0.0712	0.0349	0.0331	36	0.5239	0.0962
37	0.0000	0.0000	0.0000	0.0000	38	0.4541	0.1057
39	0.0000	0.0000	0.0000	0.0000	40	0.4541	0.1057
41	0.0000	0.0000	0.0000	0.0000	42	0.4541	0.1057
43	0.0909	0.0905	0.0378	0.0356	44	0.4541	0.1057
45	0.0000	0.0000	0.0000	0.0000	46	0.3784	0.1120
47	0.0000	0.0000	0.0000	0.0000	48	0.3784	0.1120
49	0.0000	0.0000	0.0000	0.0000	50	0.3784	0.1120
51	0.0000	0.0000	0.0000	0.0000	52	0.3784	0.1120
53	0.0000	0.0000	0.0000	0.0000	54	0.3784	0.1120
55	0.0000	0.0000	0.0000	0.0000	56	0.3784	0.1120
57	0.0000	0.0000	0.0000	0.0000	58	0.3784	0.1120
59	0.0000	0.0000	0.0000	0.0000	60	0.3784	0.1120
61	0.0000	0.0000	0.0000	0.0000	62	0.3784	0.1120
63	0.3333	0.3143	0.0946	0.0725	64	0.3784	0.1120
65	0.0000	0.0000	0.0000	0.0000	66	0.1892	0.1450
67	0.0000	0.0000	0.0000	0.0000	68	0.1892	0.1450
69	0.0000	0.0000	0.0000	0.0000	70	0.1892	0.1450

Table 2: Test of equality using log rank

Variable	df	χ^2	*p*-value
Gender	1	0.71	0.3991
Mstatus	3	4.36	0.2254
Religion	2	6.37	0.0414
WHO	3	1.62	0.6555
Disclosure	1	0.39	0.5313
Regimen	2	0.31	0.8559

Table 3: Cox proportional hazard regression model for co-infection data

Variable	df	β	SE	χ^2	*p*-value	Exp(β)
Gender						
Female	1	-0.84804	0.56113	2.2840	0.1307	0.428
Age	1	0.01125	0.02325	0.2341	0.6285	1.011
Religion compared with Traditionalists						
Christian	1	-0.34243	0.64531	0.2816	0.5957	0.710
Islam	1	-0.14885	0.71045	0.0439	0.8340	0.862
Marital status compared with widowed						
Divorced	1	0.75375	0.87429	0.7433	0.3886	2.125
Married	1	-0.04394	0.68757	0.0041	0.9490	0.957
Single	1	-3.54407	2.04189	3.0126	0.0826	0.029
Weight	1	-0.09141	0.03393	7.2594	0.0071	0.913
CD4	1	-0.00694	0.00224	9.5732	0.0020	0.993
Regimen type compared with Combivir/NVP						
AZT/3TC/EFV	1	0.10892	0.73612	0.0219	0.8824	1.115
AZT/3TC/NVP	1	-0.10173	0.65821	0.0239	0.8772	0.903
WHO clinical stage compared with IV						
I	1	0.12640	0.65050	0.0378	0.8459	1.135
II	1	0.52449	0.78025	0.4519	0.5014	1.690
III	1	-0.45564	0.75948	0.3599	0.5486	0.634
Disclosure						
No	1	0.49100	0.58050	0.7154	0.3977	1.634

Table 4: Test of proportional hazards assumption

Time	rho	χ^2	df	*p*-value
Gender	0.13198	0.67	1	0.4141
Age	-0.03238	0.02	1	0.8901
Religion	0.16932	0.45	1	0.5037
Mstatus	0.10317	0.29	1	0.5884
Weight	0.22947	1.43	1	0.2316
CD4	0.11258	1.29	1	0.2567
Regimen	0.10085	0.34	1	0.5592
WHO	0.22412	1.18	1	0.2783
Disclosure	-0.14000	1.02	1	0.3128
global test		12.00	9	0.2130

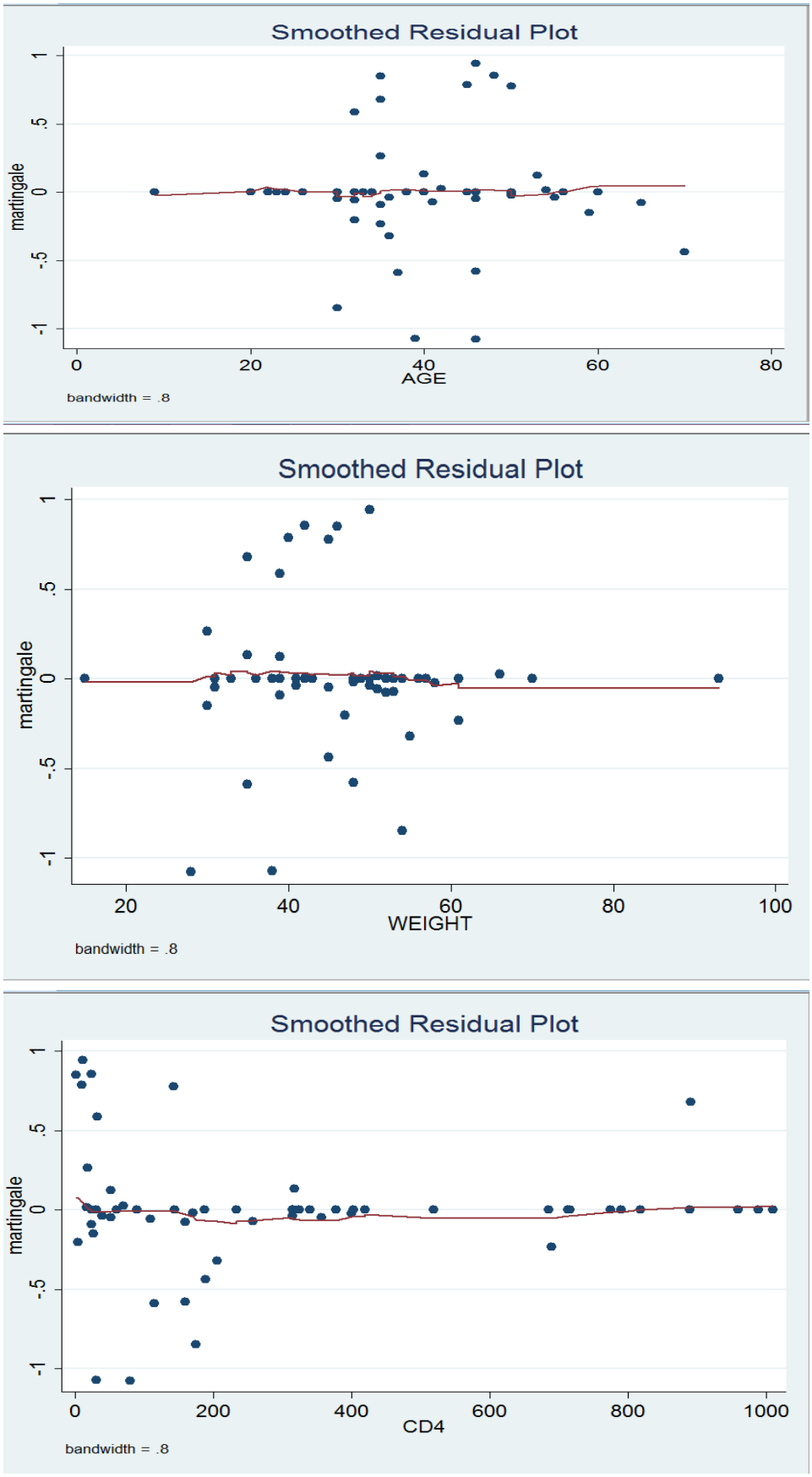

Figure 2: Martingale residual plot for continuous covariates

Table 5: Model comparison

Criterion	Weibull	Exponential	Gamma	Llogistic	Lnormal	Cox model
AIC	137.066	136.345	128.962	137.406	138.410	162.184
BIC	176.688	173.637	170.916	177.028	178.032	180.468

Table 6: Gamma model for Co-infection patients

Variable	df	β	SE	95% C.I		χ^2	*p*-value
Intercept	1	-1.5663	1.3328	-4.1785	1.0459	1.38	0.2399
GENDER							
FEMALE	1	0.1896	0.4136	-0.6209	1.0002	0.21	0.6465
Religion compared with Traditionalists							
CHRISTIAN	1	1.3132	0.5721	0.1920	2.4344	5.27	0.0217
ISLAM	1	1.3118	0.5113	0.3096	2.3140	6.58	0.0103
Marital status compared with widowed							
Divorced	1	0.3388	0.4999	-0.6409	1.3185	0.46	0.4979
Married	1	-0.3664	0.4826	-1.3123	0.5795	0.58	0.4477
Single	1	1.1728	0.8236	-0.4415	2.7871	2.03	0.1545
Weight	1	0.0809	0.0217	0.0383	0.1235	13.85	0.0002
CD4	1	0.0043	0.0007	0.0029	0.0056	36.12	0.0001
Regimen type compared with CBV/NVP							
AZT/3TC/EFV	1	-0.1321	0.5415	-1.1934	0.9292	0.06	0.8073
AZT/3TC/NVP	1	0.0747	0.3916	-0.6929	0.8422	0.04	0.8487
WHO clinical stage compared with IV							
I	1	-0.6574	0.4447	-1.5290	0.2142	2.19	0.1393
II	1	-0.3376	0.4639	-1.2469	0.5718	0.53	0.4669
III	1	-0.5659	0.5295	-1.6037	0.4719	1.14	0.2852
Disclosure							
No	1	-0.0312	0.5751	-1.1583	1.0959	0.00	0.9568
Scale	1	0.6596	0.2525	0.3114	1.3969		
Shape	1	-2.1927	1.0823	-4.3139	-0.0715		

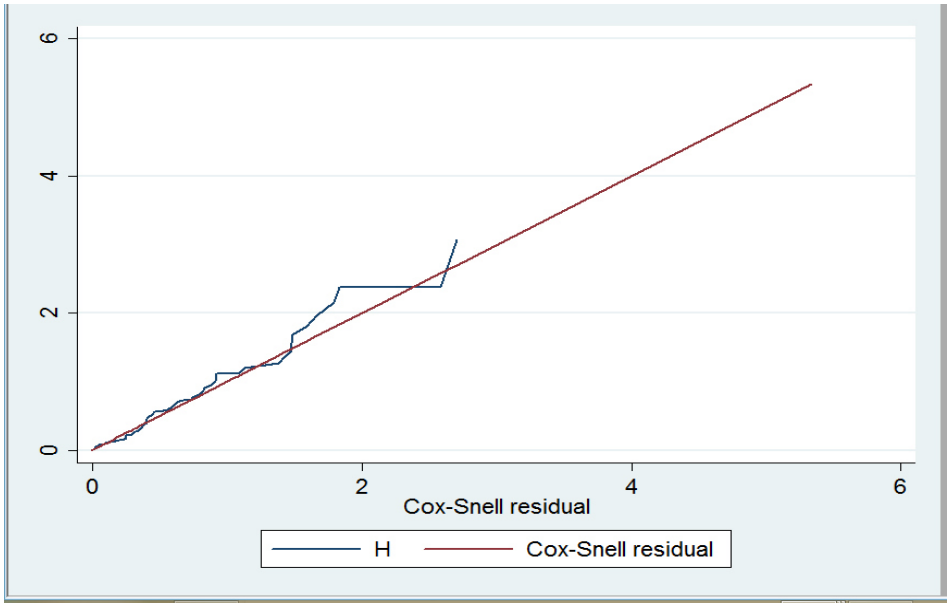

Figure 3: Cox-Snell residual plot

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar