Modeling of Tuberculosis and Tuberculosis Co-infected with Human Immunodeficiency Virus Patents using some Parametric Survival Models

A. Olaosebikan, B.Sc.[1]; S.A. Aderoju, Ph.D.[1*] and O.S. Balogun, Ph.D.[2]

[1]Department of Statistics and Mathematical Sciences, Kwara State University, Malete, PMB 1530, Ilorin, Kwara State, Nigeria.
[2]Departmet of Statistics and Operations Research, Modibbo Adamawa University of Technology, Yola, Adamawa State, Nigeria.

E-mail: samuel.aderoju@kwasu.edu.ng*
aishatolaosebikan@gmail.com
balogun.os@mautech.edu.ng

ABSTRACT

In this study, the authors examined age and gender as some of the challenges of modeling infectious diseases using data comprising of patients with Tuberculosis (TB) and TB co-infected Human Immunodeficiency Virus (HIV) as a case study. The TB and TB co-infected with HIV are some of the common health problems in the world. Time-to-event outcomes are common data types in medical research. The data examined time until a patient is cured of the disease having some patients right-censored. With the nature of the data, the appropriate analysis is the survival analysis method. The study aims at fitting appropriate models to the TB and TB/HIV co-infection data examining age and gender as factors influencing the survival of the disease. Hence, Kaplan-Meier estimation, Cox PH and some parametric models were adopted in the study.

The result shows that among the parametric models, generalized gamma fit TB data best and there is a significant difference in the survival rate of male and female while Cox model fit TB co-infected with HIV best and there is a significant difference in the male and female patient. The median survival times are 17 and 18 months for TB patients and TB/HIV co-infected patients respectively.

(Keywords: survival analysis, TB, HIV co-infection, parametric, Kaplan_Meier, Cox PH).

INTRODUCTION

An infectious disease is a disease that is caused by the invasion of a host by agents whose activities harm the host's tissue and can be transmitted to other individuals. There are six major types of infection agents, they are bacteria, viruses, fungi, protozoa, helminths, and prions (Andersion and May, 1992).

Tuberculosis is a potentially serious infectious disease that mainly affects the lungs (Mayo, 2019). The bacteria that cause tuberculosis (TB) are spread from one person to another through tiny droplets released to the air through coughs and sneezes. HIV infection is also a powerful risk factor for TB and contributes to the development of active and latent TB and exogenic re-infection (Corbett, et al., 2003).

Nigeria is among the 14 high burden countries for TB, TB/HIV and multi-drug resistant TB. It is ranked 7th among the 30 highest TB burden countries and second in Africa. The theme for 2019 World's TB Day was "It's Time" (Peterson, et al., 2019). Pathophysiology and clinical presentation of TB in children differ from what is obtained in adults. Following infectious children have a higher risk not only of progression to disease but also of extra-pulmonary dissemination and death.

Several publications have appeared in recent years documenting the works done on TB and HIV disease using different statistically analysis. Abera, et al. (2009) worked on the relationship between HIV and TB in the Oromia Regional State, Ethiopia. The study covered a total of 40,779 TB cases including 12,818 smear-positive

pulmonary TB cases and 29,950 positive of HIV infection.

The ecological association between different tuberculosis types and prevalence of HIV across zones and towns in Oromia was estimated adopting the Spearman correlation. The result of the study showed that the HIV infection prevalence was significantly associated with the incidence of TB in Oromia region $\left(r = 0.69, p < 0.01\right)$. It also showed that similar associations were noticed between HIV infection prevalence and the incidence of smear-positive tuberculosis and smear-negative tuberculosis as well as Extra-Pulmonary Tuberculosis.

Kapata, et al. (2013) showed that there was a need to explore the social determinant of TB and their association with TB/HIV co-infection was important to be addressed to have a maximum impact in the control of TB. The study had its limitations since it was conducted in an urban setting only.

Straetemans, et al. (2010) carried out a meta-analysis of cohort studies by selecting relevant articles. The purpose was to assess the effect of TB on mortality in people living with HIV. They pooled overall analysis of fifteen studies estimating the effect of tuberculosis on mortality in people living with HIV (PLWHIV) which show a Hazard ratio $\left(HR\right)$ of $2.6\ (95\%\ CI: 1.8, 3.6)$ that indicated the impact of TB on HIV in co-infection.

Zubairu (2009) conducted a study in Nigeria on the case files of patients with HIV/AIDs from January to December 2006 attending Aminu Kano Teaching Hospital. In the study, Chi-square analysis was used to test the significance of association among the categorical variable: All the variables were significantly associated with TB/HIV co-infection and were included in a multiple logistic regression analysis to determine their individual effects.

Mashimlaye (2010) considered TB treatment outcomes in adult TB patients receiving treatment at Rixile HIV clinic in Tinswalo hospital in Bushbuckridge, South Africa. The univariate analysis revealed that for the associations of age, sex, Body Mass Index (BMI), education and Antiretroviral (ARV) treatment, only age, sex and ARV treatment were discovered to predict mortality related to TB.

Refera (2012) conducted a study on the survival/death status of HIV/TB co-infected patients who were treated of TB at Ambo hospital from September 1st, 2006 to August 31st, 2011. Cox proportional hazard model covariates used significantly influence the survival of PLWHIV co-infected patients are identified.

Fatmawati (2016) proposed an optimal control on the treatment of the transmission of tuberculosis-HIV coinfection model. The optimality system was solved numerically to illustrate the effectiveness of the treatments.

Gesesew, et al. (2016) indicated that one-fifth of TB/HIV co-infected patients were deceased, and social factors seemed to have significant influence. Fatmawati (2017) presented a mathematical model on the spread of HIV and tuberculosis (TB) co-infection using the resistance of HIV to antiretroviral (ARV) drugs. The numerical simulations of the optimal control were also performed to illustrate the results obtained.

According to the study carried out by Janida, et al. (2015) reported that Abuja, Nigeria had one of the highest proportions of TB-HIV co-infection rates in Sub-Saharan Africa and it was also revealed that the outcome of patients had statistically significantly higher mortality.

Shobowale, et al. (2015) determined the demographic characteristics of patients with TB and the rate of TB/HIV co-infection from a total of 100 patient's records retrospectively and analyzed for over a 3-month period. They recommended that Tuberculosis remain a huge public health threat in Nigeria with attendant challenges in diagnosis and treatment and that improved system for the accurate diagnosis of Tuberculosis must be employed and treatment strategies are improved on and intensified.

Yu, et al. (2004) utilized a mixture cure rate model on group survival data and they observed that the estimate of the cure fraction could be quite sensitive to the length of follow up time and the choice of latency distribution (failure time distribution). They also investigated the effects of various parametric distributions such as the lognormal, log-logistic, Weibull and generalized gamma and they concluded that the estimate of the cure fraction was robust with the generalized gamma distribution. It was suggested that the accuracy of the estimate of the cure fraction is

affected when the follow-up a time long concerning the median survival time and homogeneity of the observations.

Elfaki, et al. (2013) presented a simple modification of estimating for partly-interval censored data using the semi-parametric Cox's proportional hazards regression models of the sub-distribution of a two competing risks models.

Most of the authors reviewed in this research used different methods in analyzing the disease considered but Balogun and Jolayemi (2017) came up with a generalized form of cure rate which can handle an infectious disease with possible co-infection using exponential distribution as the baseline distribution. The authors extended work by Chen (2016) which uses a Bounded Cumulative Hazard model which is a non-parametric model. They also used simulated data to demonstrate how the model works.

This research serves as a motivation for us to incorporate their idea by applying a survival analysis for a real-life data comprising of patients with TB and TB co-infected with HIV to determine the best model and estimate the risk factors associated with the diseases.

This research intends to demonstrate how survival analysis can be applied to an infectious disease with possible co-infections. The authors discovered survival analysis has been applied to infectious disease but not to a disease with possible co-infections.

The authors intend to come up with the best model to be used for the disease by fitting different parametric models to the data collected and they also intend to make some recommendations at the end of the research. The data used for this research gives little information about the patients and the analysis is restricted to the use of parametric, semi-parametric and non-parametric models.

OVERVIEW OF SURVIVAL METHODS AND THEIR APPLICATION IN MEDICAL RESEARCH

In every day of life, we want to know the time it will take for a person to recover from a particular disease, or time until the death of an individual infected with a disease (Kleinbaum and Klein, 2012). This type of situation is called survival analysis method.

Generally, the survival method is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. The time could be in days, weeks, months or years from the starting point of the follow-up of a person until an event happens; on the other hand, time can imply the age of a person when an event occurs. The event means death, disease incidence, relapse from remission, recovery (e.g., cure) or any assigned experience of intrigue that may happen to a person (Kleinbaum and Klein, 2012).

In survival analysis, we, for the most part, allude to the time variable as survival time, since it gives them time that an individual has "survived" over some follow-up period. We additionally regularly allude to the event as a failure because the event of interest often is death, disease incidence, or some other negative individual experience. In any case, survival time might be "time to cure after a treatment procedure," in which case a failure is a positive event (as the case is in this study).

Survival methods consider a key analytical issue called censoring. Generally, censoring happens when we have some data about individual survival time, however, we don't know the survival time precisely. If for a given patient, the study ends while the patient doesn't get the event, at that point that the patient's survival time is viewed as censored (Kleinbaum and Klein, 2012). We realize that, for this individual, the survival time is in any event as long as the period that the individual has been followed, however if the individual encounters the event after the study ends, we do not have the foggiest idea about the total survival time. Right-censored data can happen when an individual's actual survival time is greater than the individual's observed survival time (Kleinbaum and Klein, 2012). Left-censored data can happen when an individual's actual survival time is less than or equal to that individual's observed survival time.

DEFINITION:

The random variable,
$$T = survival\ time\ (T \geq 0)$$

Survival function, $\quad S(t) = P(T > t) \qquad$ 1

$$\text{Hazard function,} \quad h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad 2$$

$$h(t) = \frac{f(t)}{S(t)} = -\left[\frac{dS(t)/dt}{S(t)}\right]$$

$$S(t) = \exp\left[-\int_0^t h(u)du\right]$$

When either of the survival function or hazard function is known, one can get the other (Kleinbaum and Klein, 2012).

The Cox Proportional Hazards (PH) Model

A semi-parametric model is one whose functional form is unspecified. One of the most popular semi-parametric models is the cox proportional hazard model often called the Cox PH model which was proposed by David Cox (Cox, 1972). The cox model formula gives an expression for the hazard at time t for an individual with a given specification set of explanatory variables represented by X.

$$h(t, X) = h_0(t)e^{\sum_{i=1}^{p} \beta_i X_i} \qquad 3$$
$$\text{where } X = (X_1, X_1, X_1, \ldots, X_p) \to \text{explanatory variables}$$
$$h_0(t) \text{ is the baseline hazard involving t but not X's}$$

The PH assumption is that the baseline hazard is a function of t (time-dependent) but does not involve X's. The PH assumption is checked using three methods namely; graphical, the goodness of fit and time-dependent variable approaches (Kleinbaum and Klein, 2012).

In graphical method estimated lo-log curves are plotted for individuals on the same graph and it assesses the PH assumption because it shows if the PH model will be appropriate for a given set of predictors then the empirical plot log-log survival curve for different individuals is expected to be approximately parallel.

The goodness of fit for each predictor in the model, Schoenfeld residual is defined for every subject with an event. the concept is a statistical test which shows that if the PH assumption holds for a covariate then the Schoenfeld residual for such covariate will not be related to survival time

and time-dependent covariate, the effects is captured using a time-dependent function of the covariates. The proportional cox is fitted and the statistical significance of the coefficient δ can be checked (Kleinbaum and Klein, 2012).

The Cox model is widely used because it is a "robust" model. Cox is widely popular because it does not rely on the distributional assumptions for the outcome (Kleinbaum and Klein, 2012).

Weibull Model

A parametric survival model is one in which the survival time is assumed to follow a known distribution (e.g., Weibull). The time follows a probability density function, $f(t)$ and the survival and hazard function can be determined. Many parametric models are accelerated failure time (AFT) models rather than PH models.

The Weibull is a two-parameter distribution with λ and P as proposed by Weibull (Weibull, 1951). It is the most widely used parametric survival model. P is the shape parameter and determines the shape of the hazard function. The Weibull reduces to exponential when $P = 1$. The assumption AFT and PH holds for Weibull.

$$f(t) = \lambda p t^{p-1} e^{-\lambda t^p} h_0(t) \qquad 4$$

$$S(t) = e^{-\lambda t^p}$$

$$h(t) = \lambda p t^{p-1}$$

$$\text{where } \lambda = \exp(\beta_0 + \beta_1 age + \beta_2 gender)$$

The Exponential Model

The exponential is a one-parameter distribution with a constant hazard λ. The exponential satisfies both the PH and AFT assumption: (Kleinbaum and Klein, 2012).

$$f(t) = \lambda e^{-\lambda t} \qquad 5$$
$$S(t) = e^{-\lambda t}$$
$$h(t) = \lambda$$

Log-Logistic Model

The log-logistic is a two-parameter distribution λ and P. The log-logistic satisfies the AFT but not the PH model (Kaplan and Meir, 1958). The AFT model is a proportional odds model (PO) i.e. the survival odds ratio is constant over time; (Kleinbaum and Klein, 2012).

$$S(t) = \frac{1}{1 + \lambda t^p} \qquad 6$$

$$h(t) = \frac{\lambda p t^{p-1}}{1 + \lambda t^p} \qquad 7$$

$$recall, \qquad f(t) = h(t)S(t)$$

Log-Normal Model

The log-normal distribution is denoted Log-normal has 2 parameters (Galton, 1879). The shape of the distribution is similar to the log-logistic distribution and yields similar model results. The log-normal accommodates the AFT but not the proportional odds model (Aitchison and Brown, 1957) :

$$f(t) = \frac{\emptyset\left(\frac{\log(t) - \mu}{\sigma}\right)}{t\sigma} \qquad 8$$

$$F(t) = \phi\left(\frac{\log(t) - \mu}{\sigma}\right)$$

$$h(t) = \frac{f(t)}{S(t)}$$

Gamma Model

Gamma is a parametric model with two parameters used to model survival data. Gamma is an extension of the factorial function with its argument shifted down by 1 to a real and complex number. It can also be used to model service time, the lifetime of objects and repair time. The gamma distribution has an exponential right-hand tail. The gamma distribution with parameters λ and k, denoted as $\Gamma(\lambda, k)$, has probability density function:

$$f(t) = \frac{\lambda(\lambda t)^{k-1} e^{-\lambda t}}{\Gamma(k)} \qquad 9$$

$$S(t) = 1 - I_k(\lambda t),$$

where $I_k(z)$ is the incomplete gamma function, defined as

$$I_k(z) = \int_0^z \lambda^{k-1} e^{-z} \, dz / \Gamma(k) \qquad 10$$

Generalized Gamma Model

This is a parametric survival model with three parameters allowing for great flexibility in its shape. The Weibull and log-normal distribution are special cases of the generalized gamma distribution. The S(t) and h(t) are expressed in the form of integrals. A log-normal is not a proportional odds model although it satisfies the AFT model. One advantage of the parametric model compared to the cox model is that the parametric likelihood easily accommodates right, left and interval-censored data unlike the cox likelihood that easily handles right-censored data but does not directly accommodate left or interval-censored data. Binary regression is used for interval-censored data and discrete survival analysis (Lienhard and Meher, 1967). The density of the distribution can be written as:

$$f(t) = \frac{\lambda p (\lambda t)^{pk-1} e^{-(\lambda t)^p}}{\Gamma(k)}, \qquad where \ p = \sigma^{-1} \qquad 11$$

The generalized gamma model has the following as its special cases:
when p =1, we have a Gamma model

when k =1, we have a Weibull model

when p = k = 1, we have Exponential model and

when $k \to \infty$, we have a Log-Normal model

Akaike's Information Criterion

The Akaike Information Criterion (AIC) is an estimator of out-of-sample prediction error and thereby the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models.

APPLICATION TO MEDICAL RESEARCH

The data used for this work is secondary data collected by one of the authors for 15 years from 2000-2015 from the Medical Record Department of University of Ilorin Teaching Hospital (UITH), Kwara State [18]. The data comprises of patients with TB and TB co-infected with HIV with their respective age (years), gender (male and female), cure status (cured and not cured), length of stay (months) and censoring rate (censored and not censored). The data collected is made up of 518 observations of Tuberculosis and 133 observations of Tuberculosis co-infected with HIV patients. The analysis was done using "survival" (Borgan, 2001) and "flexsurv" (Jackson, 2016) packages in R statistical software (R Core Team, 2018) and STATA version 16 (StataCorp, 2019).

Ethical Consideration

Permission to implement this research was collected from the Medical Record unit of the University of Ilorin Teaching Hospital (U.I.T.H), Ilorin, Kwara State between 2000 and 2015 on patients with TB and TB/HIV co-infection. Data collected from the hospital in support of this analysis commenced immediately after the Ethics board of the University approved it.

RESULTS

The results consist of 298 males and 220 females for the TB patients considered and the average age of the patients to be 38 years. There are 53 males and 80 females for the TB co-infected with HIV patients and the age of the patients to be 37 years.
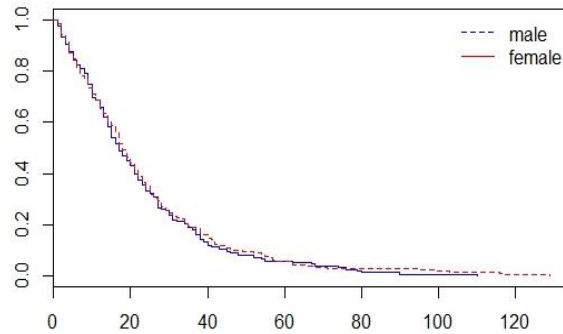
Figure 1: Kaplan Meir Curve of TB Data for Gender.

This Figure shows the survival rate of the patients with TB based on gender as it is related to male and female patient's information used.

Table 1 presents the risk factors concerning the models used and with the of the help AIC, we were able to determine the best model to be used for the TB patient's data.

Table 2 presents the risk factors in relation to all the models used and using AIC, we were able to determine the best model to be used for the TB co-infected with HIV patient's data.

Figure 2 displays the Kaplan Meir curve for the TB/HIV co-infection data.

The Cox regression model for the TB co-infected with HIV is expressed as:

$$\hat{h}(t) = \hat{h}_o(t)e^{0.02713age - 0.47920gender}$$

Table 1: Semi–Parametric and Parametric Estimate of TB Data.

	Cox PH	Expo	Weibull	L.Norm	L.Logis	Gomprtz	Gamma	Gen.G	Gen.F
Age	0.00246	0.00197	-0.00235	-0.00305	-0.0031	0.0024	0.0024	-0.0025	-0.0025
	(0.349)	(0.449)	(0.274)	(0.214)	(0.184)	(0.394)	(0.267)	(0.256)	(0.263)
Gend	-0.0621	-0.0503	0.0558	0.0424	0.0315	-0.0660	-0.0508	0.0461	0.0446
(male)	(0.518)	(0.599)	(0.480)	(0.646)	(0.724)	(0.491)	(0.526)	(0.578)	(0.594)
LL	-2334.0	-1857.9	-1844.6	-1856.7	-1855.2	-1854.3	-1841.8	-1840.7	-1840.7
DF	4	3	4	4	4	4	4	5	6
AIC	4652.09	3721.74	3697.26	3721.47	3718.40	3716.64	3691.63	3691.38	3693.36

Table 2: Semi-Parametric and Parametric Estimate of the TB/HIV Co-Infection Data.

	Cox PH	Expo	Weibull	L.Norm	L.Logis	Gomprtz	Gamma	Gen.G	Gen.F
Age	0.0271 (0.002)	0.0181 (0.032)	-0.0193 (0.000)	-0.0151 (0.025)	-0.0130 (0.045)	0.0287 (0.001)	0.0177 (0.002)	-0.0177 (0.003)	-0.0177 (0.003)
Gend (male)	-0.4792 (0.020)	-0.3425 (0.079)	0.3467 (0.006)	0.3505 (0.021)	0.3392 (0.022)	-0.5645 (0.007)	-0.3414 (0.011)	0.3415 (0.011)	0.342 (0.011)
LL	-0.520.7	-454.8	-440.3	-443.3	-443.0	-447.4	-439.8	-439.8	-439.8
DF	4	3	4	4	4	4	4	5	6
AIC	855.33	915.59	888.55	894.60	894.08	902.81	887.57	889.57	891.57
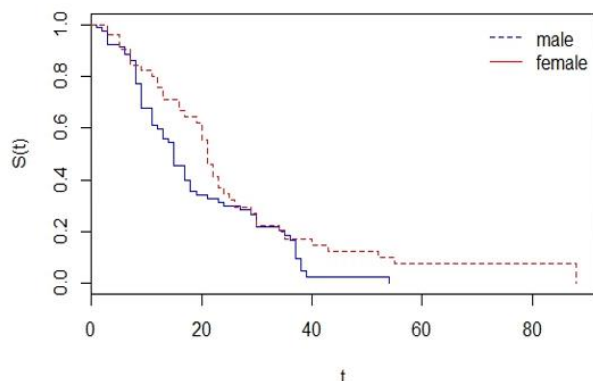
Figure 2: Kaplan Meir Curve for the TB/HIV Co-Infection Data.

DISCUSSION

The purpose of this research is to come up with the best model and examine the risk factors associated with the disease considered. The originality of the solution lies in the fact that survival analysis was applied to infectious disease with possible co-infections. And for this study, we considered TB and TB co-infected with HIV.

The Kaplan Meir curve shows that there is no significant difference in the survival time of males and females for TB patients while there is a significant difference in the survival time of males and females for TB co-infected with HIV.

The best model was selected based on the distribution with least AIC, Several parametric models were fitted to the data, in this case, we have Exponential, Weibull, Log-Normal, Log-Logistic, Gompertz, Gamma, Generalized Gamma, and Generalized-F and the semi-parametric model fitted. The Generalized Gamma fits for the TB data (AIC=3691.68). Besides, the generalized gamma showed age and gender as a significant risk factor of TB ($p < 0.05$).

In the case of TB co-infected with HIV data, the cox model fits for it (AIC = 855.367) and it shows that age and gender are significant risk factors for the data ($p < 0.05$).

CONCLUDING REMARKS

Outcome variable such as the time until an event occurs is very common in medical research. This type of variable is better to handle with survival analysis methods. In this case cure rate of tuberculosis with/without co-infection. The results of the semi-parametric Cox proportional hazard model and some selected parametric models were obtained. A key reason why the Cox model is widely popular is that it does not rely on the distributional assumptions for the outcome and robust.

In the TB data, Generalized Gamma model yields the best model. The median survival time is 17 months. It was equally shown that all the covariate was statistically significant in determining the cure rate of the disease. The Cox model outperforms other models for TB co-infected with HIV. Moreover, both covariates (that is, age and gender) are statistically significant. The median survival time is 18 months. The survival rates for females with TB-HIV co-infection are higher than that of males across age.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

REFERENCES

1. Abera, B., F. Kate, H. Zelalam, and F. Andrew. 2009. "The Association of TB with HIV Infection in

Oromia Regional National State, Ethiopia in 2006/7". *Journal of Ethiopia Health Dev.* 23(1): 63-67.

2. Aitchison J. and J.A.C. Brown. 1957. *The Log-Normal Distribution*. Cambridge University Press: London, UK.

3. Anderson R.M and R.M. May. 1992. *Infectious Disease of Humans: Dynamics and Control.* Oxford University Press: New York, NY.

4. Balogun, O.S. and E.T. Jolayemi. 2017. "Modelling Cure Rate Model for TB with HIV Co-infection". *Pacific Journal of Science and Technology*. 18(1):288-299.

5. Borgan, O. 2001. "Modeling Survival Data: Extending the Cox Model". In: *Statistics in Medicine*. M.T. Therneau and P.M. Grambsch (ed.). Springer-Verlag: New York, NY. 350. ISBN 0-387-98784-3.

6. Chen, Tai-Tsang. 2016. "Predicting Analysis Times in Randomized Clinical Trials with Cancer". Petersen, E., et al., 2019. World Tuberculosis Day March 24th 2019 Theme:"It's TIME"—*International Journal of Infectious Diseases Tuberculosis Theme Series. International Journal of Infectious Diseases*, 80, S1-S5.

7. Corbett, E.L, C.J. Watt, N. Walker, D. Maher, B.G. Williams, M.C. Raviglione, and C. Dye. 2003. "The Growing Burden of Tuberculosis: Global Trend and Interactions with the HIV Epidemic". *Arch Intern Med.* 163(9):1009-21.

8. Cox, D.R. 1972. "Regression Models and Lifetables". *J. Royal Stati Society*. 34: 187-220.

9. Elfaki, F.A.M., A. Abobakar, M. Azram, and M. Usman. 2013. "Survival Model for Partly Interval-Censored Data with Application to Anti D in Rhesus D Negative Studies". *International Scholarly and Scientific Research and Innovation*. 7(5): 895-898.

10. Fatmawati, H.T. 2016. "An Optimal Treatment Control of TB-HIV Co-infection". *International Journal of Mathematics and Mathematical Sciences*. ID8261208.

11. Fatmawati, H.T. 2017. "Optimal Control of HIV Resistance and Tuberculosis Co-infection using Treatment Intervention", *Asian Pacific Journal of Tropical Disease*. 7(6): 366-373.

12. Galton, F. 1879. "The Geometric Mean in Vital and Social Statistics". *Proceedings of the Royal Society of London.* 29: 365–7.

13. Gesesew, H., B. Tsehaineh, D. Massa, A. Tesfay, H. Kahsay, and L. Mwanri. 2016. "The Role of Social Determinants on Tuberculosis/HIV Co-infection Mortality in Southwest Ethopia: A Retrospective Cohort Study". *BMC Research Notes.* 9(1): 89.

14. Jackson, C. 2016. "Flexsurv: A Platform for Parametric Survival Modeling". *R. Journal of Statistical Software.* 70(8): 1-33. doi:10.18637/jss.v070.i08

15. Janida, M.A., L. Lawson, O.O. Nnodu, M.N. Ajani, E.O. Umobong, C.C. Fredrick, R. Adelaiye, W.D. Nafarda, S.T. Abdulraham, J.N. Okechukwu, I.T. Tafida, I. Bashir, N.D. Baamlong, R.Q. Gurgel, M.A. Yassin, O. Oladimeji, and L.E. Cuevas. 2015. "Treatment Outcome of Patients Co-infected with Tuberculosis and HIV in Abuja, Nigeria". *Nigerian Journal of Basic and Clinical Sciences*. 11(2): 72-75.

16. Kapata, N., P. Chanda-Kapata, and C. Michelo. 2013. "The Social Determinants of Tuberculosis and their Association with TB/HIV Co-infection in Lusaka, Zambia". *Medical Journal of Zambia*. 40(2): 48-54.

17. Kaplan, E.L. and P. Meir. 1958. "Non-Parametric Estimation from Incomplete Observations". *Journal of the American Statistical Association*. 53(232): 457-481.

18. Kleinbaum, D.G. and M. Klein. 2012. "Survival Analysis; A Self-Learning Text". *Third Edition, Statistics for Biology and Health, Springer Science + Business Media*. ISBN 9781441966452 .

19. Lienhard J.H and P.L. Meyer. 1967. "A Physical Basis for the Generalized Gamma". *Quarterly of Applied Mathematics*. XXV:3.

20. Mashimbye L. 2010. "Tuberculosis (TB) Treatment Outcomes in Adult TB Patients Attending a Rural HIV Clinic in South Africa (Bushbuckridge). Available at: http://hdl.handle.net/10539/7982.

21. Mayo Clinic. 2019. "Tuberculosis - Symptoms and Causes". Mayo Foundation for Medical Education and Research (MFMER): Rochester, MN. https://www.mayoclinic.org/diseases-conditions/ tuberculosis/symptoms-causes/syc20351250

22. Petersen, E., M. Rao, G. Ippolito, G. Gualano, J. Chakaya, F. Ntoumi, D. Moore, R. Allen, K. Gaskell, J.N. Öhd, and M.P. Hergens. 2019. "World Tuberculosis Day March 24th 2019 Theme: "It's TIME". International Journal of Infectious Diseases Tuberculosis Theme Series. *International Journal of Infectious Diseases*. 80, S1-S5.

23. R Core Team. 2018. "R: A Language and Environment for Statistical Computing". *R*

Foundation for Statistical Computing: Vienna, Austria. *https://www.R-project.org/*.

24. Refera, H. 2012. "Survival and Risk Factors of HIV/TB Co-infected Patients under Antiretroviral Therapy in Ambo Hospital, Ethiopia". Master's Thesis (unpublished), Department of Statistics, Addis Ababa University.

25. Shobowale, O., C.J. Elikwu, and B. Adegunle. 2015. "Tuberculosis and HIV Co-infection in a Tertiary Hospital in South Western Nigeria". *International Journal of Medical and Biomedical Science*. 3(1): 1-4.

26. StataCorp. 2019. "Stata Statistical Software: Release 16". StataCorp LLC.: College Station, TX.

27. Straetemans, M., A.L. Bierranbach, N. Nagelkerke, P. Glaziou, and M.J. van der Weff. 2010. "The Effect of Tuberculosis on Mortality in HIV Positive People: A Meta-Analysis". *PLoS ONE*. 5(12): 1-10.

28. Weibull, W. 1951. "A Statistical Distribution Function of Wide Applicability". *Journal of Applied Mechanics*. 18:293-297.

29. Yu, B., R.C. Tiwari, K.Z. Cronin, and E.J. Feuer. 2004. "Cure Fraction Estimation from the Mixture Cure Rate Models for Grouped Survival Data". *Statistics in Medicine*. 23: 1733-1747.

30. Zubairu, L. and B. Musa. 2009. "Prevalence and Predictors of Tuberculosis Co-infection among HIV-Seropositive Patients Attending the Aminu Kano Teaching Hospital, Northern Nigeria". *J. Epidemiol*. 19(2): 81-87.

ABOUT THE AUTHORS

A. Olaosebikan, is a B.Sc. Statistics graduate student of Department of Statistics and Mathematical Sciences, Kwara State University, Malete, Kwara State, Nigeria. Her research interest is survival analysis.

S.A. Aderoju, is a Lecturer at the Department of Statistics and Mathematical Sciences, Kwara State University, Malete, Kwara State, Nigeria. He holds a Ph.D. in Statistics from the Department of Statistics, University of Ilorin, Kwara State, Nigeria. He is a member of Nigeria Statistical Association (NSA). He has published several articles in national and international bodies. His research interest is survival analysis, biostatistics and probability distribution models.

O.S. Balogun, is a Lecturer at the Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Adamawa State, Nigeria, and he is currently a Post-Doctoral Researcher at the School of Computing, University of Eastern Finland. He holds a Ph.D. in Statistics from the Department of Statistics, University of Ilorin, Kwara State, Nigeria. He is a Senior Research Fellow at Centre for Multidisciplinary Research and Innovation (CEMRI) and He is a member of International Biometric Society (IBS), International Statistical Institute (ISI), Nigerian Statistical Association (NSA), Nigerian Mathematical Society (NMS). He has several conference and journal papers in reputable international bodies, and he is also served as guest reviewers for several journals. His research interest is data science, machine learning, biostatistics, categorical data analysis, modeling and probability distribution models.

SUGGESTED CITATION

Pacific Journal of Science and Technology