Advances in Mathematical & Computational Sciences Journal,
Vol. 6 No. 2, June, 2018

Article Citation Format
Akingbade, A.A.; Ikegwu, E.M ; Badmus, N.I. ; Akinsola, O.J. (2018). Survival Analysis: An Application To Patients With Tuberculosis Infection. Journal of Digital Innovations & Contemp Res. In Sc., Eng & Tech. Vol. 5, No. 4. Pp 11-20.

Article Progress Time Stamps

Article Type: Research Article
Manuscript Received: 13rd May, 2018
Review Type: Blind
Final Acceptance:: 19th September, 2018
DOI Prefix: 10.22624

Survival Analysis: An Application To Patients With Tuberculosis Infection

[1]Akingbade, A.A.; [1]Ikegwu, E.M ; [1]Badmus, N.I. ; [2]Akinsola, O.J.
[1]Department of Statistics, Yaba College of Technology, Yaba, Lagos, Nigeria
[2] Department of Community and Primary care, College of Medicine,
University of Lagos, Lagos, Nigeria
Email Address: (kunle.akingbade@gmail.com)
Phone Number: (08036321288)

ABSTRACT

Tuberculosis is one of the leading causes of morbidity in developing countries and Nigeria ranks seventh among the countries with high burden of the disease in the world and second in Africa. This study employs cox regression to examine the risk factor associated with Tuberculosis infection. The variables considered are, age, gender, weight, HIV status and diabetes status. The data used for this research were obtained from the record of Patient with Tuberculosis infection at the Lagos University Teaching Hospital (LUTH), Lagos, comprising one hundred and thirty six (136) males and one hundred and twenty six (126) females. The results from the analysis show that show that there is no significant difference in the survival curves of males and females, The probability of having tuberculosis is higher in female by 8.1% than in male. The risk of having tuberculosis is lowers in older people than younger ones. The probability of having tuberculosis is higher by 55.7% in those who are tested positive for diabetes than those who are not. Hence, there is higher association between diabetes mellitus and tuberculosis. Furthermore, the risk of contracting tuberculosis is higher by 9.3% in those who are tested positive with HIV than those who are not. Hence, HIV is also a risk factor of tuberculosis. Low body weight is also associated with tuberculosis infection. Body weight, HIV status and Diabetes status are statistically significant in predicting the chance of contracting tuberculosis infection among other factors considered.

Keywords: Cox Regression, Tuberculosis, Diabetes Mellitus, Survival function, Hazard function

1. BACKGROUND TO THE STUDY

Tuberculosis remains a global public health challenge of great dimension. It happens to be the unprecedented world's most infectious deadly killer with about 4500 lives lost per day and unfortunately, Nigeria is far worse hit by this global epidemic in Africa. Nigeria currently ranks 7th in the world and 2nd in Africa among the 30 countries with the highest burden of TB, TB/HIV, multi drug resistant TB. Significant progress has been made in the fight against Tuberculosis, but it continues to be a life-threatening disease that is worsened by many challenges responsible for its prevalence.

Although curable, treatment and diagnosis for Tuberculosis continue to be a matter of global concern, especially in the emergence of multidrug-resistant TB (MDR TB) which poses a major health security threat and jeopardizes long running global efforts to curb the deadly disease Tuberculosis (TB) is a chronic infectious disease caused by bacteria generally referred to as mycobacterium tuberculosis; almost every organ in the body can be affected, but involvement of the lungs account for more than 80% of TB cases. Tuberculosis affecting the lungs is called Pulmonary Tuberculosis (PTB), while those affecting other organs are called Extra Pulmonary Tuberculosis (EPTB), The most important source of infection is an untreated Pulmonary Tuberculosis (PTB) patient. When such a person coughs, spits or sneezes, tiny droplet nuclei containing the tubercles are released. Transmission is through inhaling these droplet nuclei (Federal Ministry of Health 2010). Tuberculosis is responsible for more deaths than any other infectious disease (WHO, 2008). It was estimated to cause a global emergency with estimates of 1.8 million deaths worldwide in 2008 out of over nine million cases. In the same year, the estimated global incidence rate fell to 139 cases per 100,000 populations after reaching its peak in 2004 at 143 per 100,000. However, this decline was not homogeneous throughout the World Health Organization (WHO) regions, with Europe failing to record a substantial decline, but rather appearing to have reached a stabilization rate (WHO, 2009). In the WHO African region with a population estimate of 836,670,000 as at 2010, Nigeria ranking the tenth among the 22 high TB burden countries in the world has the prevalence of 133 per 100,000 and 93,050 cases been registered in 2010 (Federal Ministry of Health, 2011).

TB has become a challenging development problem because one of the major factors fueling its prevalence is poverty; sadly, we have about 152million Nigerians living below poverty line (WHO, 2018). WHO reports that about 2 million people die from TB yearly and 10.4 million new cases of TB were reported in 2016, with seven countries accounting for 64% of the burden comprising of India, Indonesia, China, Philippines, Pakistan, Nigeria and South Africa[34]. Sadly, many of these people affected by TB are poor and disadvantaged people who live in impoverished communities with remote access to healthcare and because TB infected persons also experience stigma and discrimination, many TB cases go untreated. Transmissions of the disease continue unabated and because TB is airborne, the effects are devastating.

2. STATEMENT OF PROBLEM

According to World Health organization, Tuberculosis happens to be the unprecedented world's most infectious deadly killer with about 4500 lives lost per day and unfortunately, Nigeria is far worse hit by this global epidemic in Africa. Nigeria currently ranks 7th in the world and 2nd in Africa among the 30 countries with the highest burden of TB, TB/HIV, multi drug resistant TB. Significant progress has been made in the fight against Tuberculosis, but it continues to be a life-threatening disease that is worsened by many challenges responsible for its prevalence. This study derives its motivation from the rate at which TB is spreading in Africa especially Nigeria and intend to examine the risk factors associated with TB infections We are motivated by this statistics and aimed at achieving the following objective

3. OBJECTIVE

The main objective of this study is to examine how survival analysis can explain the risk factors associated with outcome of Tuberculosis infection.

4. METHODOLOGY

4.1 The Research Design

We employed Cox regression model to conduct a survival analysis of patients with tuberculosis infection in Lagos University Teaching Hospital (LUTH).The Cox regression model is a statistical theory of counting processes that unifies and extends non-parametric and parametric approaches to statistical inference [24]. A parametric model based on exponential distribution is written as:

$$log h_i(t) = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} \quad (1)$$

Equivalently,

$$h_i(t) = exp^{(\alpha+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik})} \quad (2)$$

where the constant α in the model represents log-baseline hazard. Cox regression in contrast, leaves the baseline hazard function $\alpha(t) = log h_0(t)$ $unspecified$:

$$log h_i(t) = \alpha(t) + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} \quad (3)$$

Equivalently,

$$h_i(t) = h_0(t) \exp^{(\beta_1 x_{i1}+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_k x_{ik})} \quad (4)$$

Where
t, represents the survival time
$h_i(t)$, represents the hazard function determined by set of k covariate $x_1, x_2, \ldots x_k$
h_0 is the baseline hazard that corresponds to the value of hazard if all the x_i are equal to zero

The quantity $\exp(^{\beta i})$ are called the hazard ratio.

A key assumption of the Cox model is that the hazard curves for the groups of observation should be proportional and cannot cross.

Consider two observations k and k' that differ in their x values. The corresponding hazard function can be simply written as follow:

$$h_k(t) = h_0(t) e^{\sum_{i=1}^{n} \beta_i x_i} \quad (5)$$

$$h_{k'}(t) = h_0(t) e^{\sum_{i=1}^{n} \beta_i x_{i'}} \quad (6)$$

The hazard ratio of the two observations is:

$$\frac{h_k(t)}{h_{k'}(t)} = \frac{h_0(t) e^{\sum_{i=1}^{n} \beta_i x_i}}{h_0(t) e^{\sum_{i=1}^{n} \beta_i x_{i'}}} = \frac{e^{\sum_{i=1}^{n} \beta_i x_i}}{e^{\sum_{i=1}^{n} \beta_i x_{i'}}} \quad (7)$$

This is independent of t. Consequently, the cox model is proportional hazard model

4.2 Validation of Proportional Hazard Assumption

Model adequacy is necessary after the model has been fitted. The main assumption of the Cox proportional hazard model is proportional hazard which means that the ratio is constant over time. Several methods are available to verify this assumption of proportionality (Graphical method, scaled Schoenfeld residual , Adding time dependent covariate[12].

4.2.1 Graphical Method

Based on Cox regression model, the survival function for i^{th} individual is given by:

$$S_i(t) = [S_0(t)]\exp[\beta' x_i] \quad (8)$$

where $x = (x_1, x_2, \ldots x_k)$ is a vector of explanatory variables for a particular individual.
Taking double log, we get:

$$ln[-S_i(t) = \beta' x_i + ln[-S_0(t)] \quad (9)$$

The difference in log-log curves corresponding to two different individuals with variable

$x_1 = (x_{11}, x_{12}, \ldots x_{1k})$ and $x_2 = (x_{21}, x_{22}, \ldots x_{2k})$ which does not depend on t is given by:

$$ln[-S_i(t,\ x_1)] - ln[-S_i(t,\ x_2)] = \sum_{i=1}^{k} \beta(x_{1i} - x_{2i}) \quad (10)$$

This provides the basis for assessing the validity of proportional hazard assumption. By plotting estimated $-\log[-\log(survival)]$ against survival time for the two groups, we would see parallel curve if the hazards are proportional [20].

4.2.2 Scale Schoenfeld Residuals

Scale Schoenfeld residuals are defined as a product of the inverse of the estimated variance- covariance matrix of the k^{th} Schoenfeld Residual and the k^{th} Schoenfeld Residual [25]. The scaled Schoenfeld Residual can be used to assess time trends and lack of proportionality

$$r^*_{pji} = (V^{-1})\ r_{pji} \quad (11)$$

Where r^*_{pji} is the scaled Schoenfeld residual and the r_{pji} is the Schoenfeld residual. Under the null hypothesis, we expect to see a constant function over time. When the proportional assumption holds, straight horizontal line with zero slope is expected [29].

5. DATA ANALYSIS

Table 1: Omnibus Tests of Model Coefficients

-2 Log Likelihood	Overall (score)		
	Chi-square	Df	Sig.
1976.296	10.785	5	.029

The model Chi – Square derived from the likelihood of observing the actual data under the assumption that the model fitted is accurate [$\chi^2_5 = 10.218, p = .029$] .The null hypothesis that the model is not fit is therefore rejected. This indicates that the model has a good fit. This suggests that the existence of a relationship between the independent variables and the dependent variable is supported.

Table 2: Parameter Estimates of the model

	B	SE	Wald	Df	Sig.	Exp (B)	95.0% CI for Exp(B)	
							Lower	Upper
Age	-.003	.006	.241	1	.624	.997	.986	1.008
Gender	.078	.137	.323	1	.570	1.081	.826	1.415
Weight (kg)	.012	.006	4.108	1	**.043**	1.012	1.000	1.023
Diabetes Status	.443	.178	6.188	1	**.013**	1.557	1.098	2.206
HIV Status	.089	.205	.188	1	.665	1.093	.731	1.633

Table 2 show important parameters like the hazard ratio and its associated probabilities. The hazard ratio [Exp (B)] is the relative hazard corresponding to a unit change in the associated predictors (age, gender, weight, diabetes and HIV status). The hazard ratio associated with gender is 1.081 implying that the tuberculosis is 8.1% more likely in male than in female. Also, the hazard ratio associated with weight is 1.012 implying that people with low weight is 1.2% more likely to have tuberculosis infection than people with normal weight. In addition, the hazard ratio associated with age is 0.997 which implies that the probability of having tuberculosis is reduced with age by 0.3%. This implies that younger people are at greater risk of infected with tuberculosis .Furthermore, the hazard ratio associated with diabetes status is 1.557 implying that those who tested positive of diabetes are 55.7% more likely of infected with tuberculosis than those who do not. Lastly, the hazard ratio associated with HIV status is 1.093 meaning that the risk of contracting tuberculosis is higher in those who tested positive than those who do not by 9.3%.

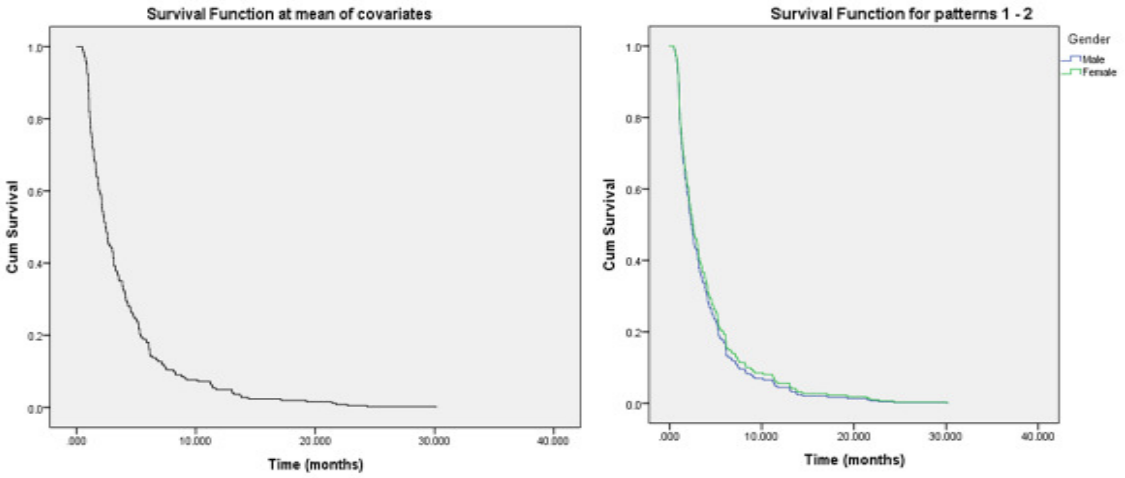

Fig. 1: Graph of Cumulative total survival and gender based survival vs. time

The survival curves in fig. 1 display the model predicted time to tuberculosis infection for the "average" patient. The horizontal axis shows the time to event. The vertical axis shows the probability of survival. It is clear from the plot that the risk of surviving tuberculosis is lower in both female and male

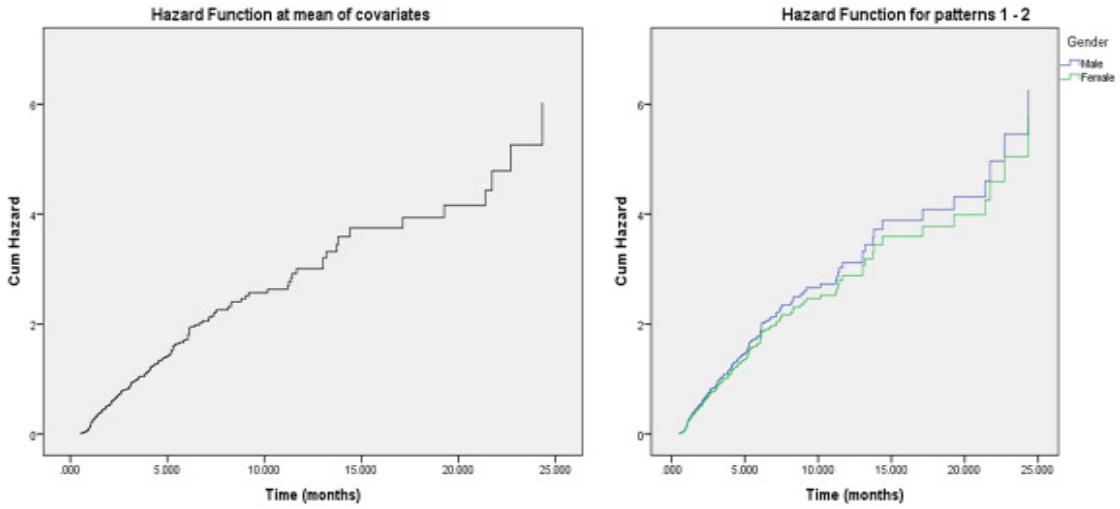

Fig. 2: Graph of cumulative hazard functions

Fig.2 shows that the risk of getting tuberculosis increases with time overall and in both male and female. Hence, the hazard rate is a function of time increases over time.

Test of proportional-Hazards Assumption

Table 3: Global Test of Proportionality Hazard Assumption

	Chi - square	df	Prob>chi2
Global test	2.42	5	0.7892

Prior to conducting a cox regression, the relevant assumptions of this statistical analysis were tested. Firstly, a sample size of 264 was deemed adequate given five independent variables to be included in the analysis. An examination of global test (see Table 3) revealed that the data under study did not violate the assumption ($p > 0.05$)

Test of proportional-Hazards Assumption (Graphically)

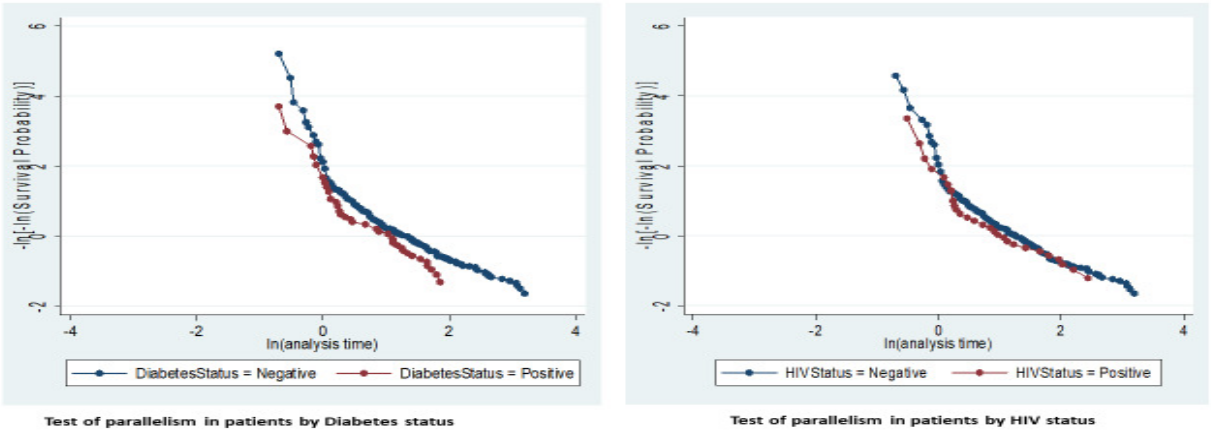

Fig. 3: Test of parallelism in patients by both Diabetes and HIV status

This is the graphical method for assessing violations of the proportional hazard assumption. Although, using graphs to assess the validity of the assumption is subjective. These are often referred to as "log – log" plots. Examining the proportional-hazards assumption on a variable without adjusting for covariates is usually adequate as a diagnostic tool before fitting Cox model. Hence, the simplest method is to check for parallel lines in the Log (-Log) plot of survival (graphically) and since the plotted lines are reasonably parallel, the assumption has not been violated (fig. 3)

6. DISCUSSION OF FINDINGS

Prior to conducting a cox regression, an examination of global test (see Table 4) and log – log plot (fig. 3) revealed that the proportional hazard assumption were not violated. This is because the global test has a Chi – square value of 2.42 with 5 degrees of freedom and the probability value above .05 while the graphical assessment shows that the plotted lines are parallel. The overall significance is tested using the model Chi – square in Table 1, which is derived from the likelihood of observing the actual data under the assumption that the model that has been fitted is accurate and showed that the model is a good fit ($\chi^2 = 10.785$, $p < 0.05$).

The effects of individual predictors (Table 2) is explained by $Exp(B)$, which is the hazard ratio and can be interpreted as the predicted change in the hazard for a unit increase in the predictor. Therefore, the result signifies the probability of having tuberculosis is higher in male by 8.1% than in female and also, that people with low body weight is at risk of tuberculosis infection by 1.2%. Hence, low body weight is associated with risk of tuberculosis.

In addition, the results also reveal that the risk of having tuberculosis is lowers in older age than younger people and that the probability of having tuberculosis is higher by 55.7% in those who are tested positive to diabetes than those who are not. Hence, there is higher association between diabetes mellitus and tuberculosis. People living with diabetes mellitus are at increased risk of contracting tuberculosis that supports the literature.

Lastly, the hazard ratio associated with HIV status show that the risk of contracting tuberculosis is 9.3% higher among people living with HIV. Hence, HIV is also a risk factor of tuberculosis and this result agreed with Ayles, et al. (2009) and Bailey & Ayles (2017) who found higher proportion of TB among HIV positive patients. And that weight and Diabetes status are statistically significant in predicting the contracting of tuberculosis infection among other factors considered. While the presence of Diabetes increases the risk of tuberculosis infection supporting previous findings (Zheng, Hu, & Gao,(2017); Workneh, Bjune, & Yimer, 2017; Oni, et al., 2017; Bailey & Ayles, 2017; Ali, Khalid, & Manju, 2016), lower weight increases the risk of tuberculosis infection corroborating the previous findings (Berkowitz, et al., 2018; Adeyeye, Ogbera, & Kuyinu, 2013)

7. CONCLUDING REMARKS

- There was no significant difference in the survival curves of male and female and among age distributions but female had a higher hazard rate of contracting tuberculosis than male.
- Weight and Diabetes status are statistically significant in predicting the likelihood of contracting of tuberculosis infection. Hence, people living with diabetes mellitus are at increased risk of contracting tuberculosis as well as those who have low body weight.
- People living with HIV are at increased risk of contracting tuberculosis.

8. CONTRIBUTIONS TO KNOWLEDGE

In order to avert the astronomical increase in people with tuberculosis infection, the following recommendations are suggested:

- Individuals should take their nutrition seriously so as to avoid low weight.
- Individual should do a regular check of their insulin level for early detection of diabetes and treatment
- Further research should be carried out that will include other risk factors of contracting tuberculosis such as smoking, alcohol , height, body mass index (BMI) etc.

REFERENCES

1. Adeyeye, O. O., Ogbera, A., & Kuyinu, Y. (2013). Prevalence of diabetes mellitus in persons with tuberculosis in a tertiary health centre in Lagos, Nigeria. Indian Journal of Endocrinology and Metabolism, 17(3), 486 - 489. doi:10.4103/2230-8210.111646
2. Akinola, A. F. , Abimbola, S. O., & Afolabi, E. B. (2009). Treatment Outcomes Among Pulmonary Tuberculosis Patients at Treatment Centres. Annals of African Medicine, 8(2), 100-104.
3. Ali, N. S., Khalid, U. K., & Manju, S. (2016). Effect of Diabetes Mellitus on Tuberculosis Treatment Outcomes and Adverse Reactions in Patients receiving Directly Observed Treatment Strategy in India: A Prospective Study. BioMed Research International [online, 7273935 (11). doi:10.1155/2016/7273935
4. Antoine, D., French, C. E., Jones, J., & Watson, J. M. (2007). Tuberculosis Treatment Outcome Monitoring in England, Wales & Northern Ireland for Cases Reported in 2001". Epidemiology of Community Health, 302 - 307.
5. Ayeles, H., Schaap, A., Nota, A. S., Sismanidis, C., Tembwe, R., De Haas, P., . . . Godfrey - Faussett, P. (2009). Prevalence of Tuberculosis, HIV and Respiratory Symptoms in two

Zambian Communities: Implications for Tuberculosis Control in the Era of HIV. PlosOne [online], 4(5), e5602.

6. Bailey, S. L., & Ayles, H. (2017). Association between diabetes mellitus an active tuberculosis in Africa and the effect of HIV. Tropical Medicine & International Health, 22(3), 261 - 268. doi:10.1111/tmi.12822
7. Berkowitz, N., Okorie, A., Goliath, R., Levith, N., Wilkinson, R. J., & Oni, T. (2018). The prevalence and determinants of active tuberculosis among diabetes patients in Cpae Town, South Africa, a high HIV/TB burden setting. Diabetess Research and Clinical Practice, 138, 16 - 25.
8. Biadglegne, F., Anagaw, B., Debebe, T. A., Tessema, B., & Rodloff, A. (2013). A Retrospective Study on the Outcome of Tuberculosis treatment in FelegeHiwot Referral Hospital. International Journal of Medical Sciences [online], 5(2).
9. Biadglegne, F., Tesfaye, W., Anagaw, B., Tessema, B., Debebe, T., Anagaw, B., Rodloff, A. C. (2013). Tuberculosis Lymphadenitis in Ethiopia. Japanese Journal of Infectious Diseases, 66(4), 263 - 268.
10. Borucka, J. (2014). Methods for Handling Tied Events in the Cox Proportional Hazard Model. Studia Oeconomica Posnaniensia, 2(2 (263)), 91 - 106.
11. Breslow, N. (1975). Covariance Analysis of Censored Survival Data. Biometrics, 30(1), 89 - 99.
12. Collet,D.(1994). Analysis of Survival data in medical research. Chapman & Hall, London.
13. Cox, D. R. (1972). Regression Models and Life Tables (with discussions). Journal of the Royal Statistical Society, Series B(34), 187 - 220.
14. Dooley, K. E., & Chaisson, R. E. (2009). Tuberculosis and diabetes mellitus: Convergence of two epidemics. Lancet Journal of Infectious Diseaes, 9(12), 737 - 746.
15. Dooley, K. E., Tang, T., Golub, J. E., Dorman, S. E., & Cronin, W. (2009). Impact of diabetes mellitus on treatment outcomes of patients with active tuberculosis. American Journal of Tropical Medicine and Hygiene, 80(4), 634 - 639.
16. Dye, C., Maher, D., Weil, D., Espinal, M., & Raviglione, M. (2006). Targets for global tuberculosis control. International Journal of Tuberculosis & Lung Diseases, 10, 460 - 462.
17. Efron, B. (1977). The Efficiency of Cox's Likelihood Function for Censored Data. Journal of the American Statistical Association, 72(356), 557 - 565.
18. Fernandez, D. K., & Manissero, D. A. (2008). Framework Action plan to fight tuberculosis in the European Union. EuroSurveillance, 18(13).
19. FMC. (2010). National Tuberculosis and Leprosy Control Programme (NTBLCP) - Workers Manual 5th. Ed. Abuja: Department of Public Health, Federal Ministry of Health.
20. Grambsch, P. and Thermeau, T.M.(1994). Proportional Hazard Test and Diagnostics Based on weighted residuals. Biometrika. 18: 515-526
21. Grill, Richard. D.(1984). Understanding Cox's regression model : A martingale Approach. Journal of American statistical association, 79:441- 447
22. Guclu, I., & Cetinkaya, N. (2015). Cox Proportional Hazards Model in Social Science. International Journal of Social Science, 36, 63 - 74.
23. Hosmer, D. W., & Lemeshow, S. (1999). Applied Survival Analysis: Regression Modelling of Time to Event Data. New York: John Wiley & Sons.
24. Klembaum, D.G.(1996). Survival: A self learning text. Spring , New york.
25. Medhat Mohamed, A.A., Sally Hossam Eldin, A.Z.(2015). Modelling survival data using Cox regression. American Journal of Theoretical and Applied statistics.4 (6), 504 – 512

26. Naidoo, P., Jinabhai, C. C., & Taylor, M. (2007). Role and Contribution of Private Healthcare Sector Doctors in the Management of HIV - infected Patients in the eThekwini Metropolitan area of KwaZulu - Natal. The SOuthern African Journal of Epidemiology and Infection, 22(22), 13 - 17. doi:10.1080/10158782.2007.11441278
27. Oni, T., Berkowitz, N., Kubjane, M., Goliath, R., Levitt, N. S., & Wilkinson, R. J. (2017). Trilateral overlap of tuberculosis, diabetes and HIV - 1 in a high burden African setting: implications for TB control. European Respiratory Journal [online], 50, 1700004.
28. Rios , M., Garcia, J., & Sanchez, J. (2000). A statistical analysis of the seasonality in pulmonary tuberculosi. European Journal of Epidemiology, 483 - 488.
29. Stel, V. S., Dekker, F. W., Tripepi, G., Zoccali, C., & Jager, K. J. (2011). Survival Analysis II: Cox Regression. Nephron Clinical Practice, 119, 255 - 260. doi:10.1159/000328916
30. Therneau, T. M., & Grambsch, P. M. (2000). Modeling Survival data, extending the cox model. New york: Springer.
31. WHO. (2008). Global Tuberculosis Control: Surveillance, Planning, Control: WHO Report 2008 . Geneva: World Health Organisation.
32. WHO. (2009). Global Tuberculosis Control 2009: Epidemiology, Strategy, Financing. Geneva: World Health Organisation.
33. WHO. (2015).Implementing the End Tb Strategy: The Essentials. http://www.who.int/tb/publications/2015/end_tb_essential.pdf?ua=1
34. WHO. (2017). Global Tuberculosis Report 2017. Geneva: World Health Organisation.
35. WHO.(2018).WorldTBDay2018,WhatCanIdo?Source *http://www.stoptb.org/events/world_tb_day/*
36. Workneh, M. H., Bjune, G. A., & Yimer, S. A. (2017). Prevalence and Associated Factors of Tuberculosis and Diabetes Mellitus Comorbididty: A Systematic Review. PlosOne [online], 12(4), e0175925. doi:10.1371/journal.pone.0175925
37. Zheng, C., Hu, M., & Gao, F. (2017). Diabetes and pulmonary tuberculosis: a global overview with special focus on the situation in Asian countries with high TB - DM burden. Global Health Action [online], 10(1264702), 1 - 11. doi:10.1080/16549716.2016.1264702
38. Zhou, M. (2001). Understanding the Cox Regession Miodels with Time - Change Covariates. The American Statistician, 55.

Acknowledgement

We acknowledge the management of Lagos University Teaching Hospital (LUTH), Lagos for the release of the data used in this study.