International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

Evaluation of Some Factors Influencing Tuberculosis Patients: A Survival Analysis

Ogunlade Temitope Olu[1,*], Akindutire Opeyemi Roselyn[2]
[1]*Department of Mathematics, Ekiti State University, Ado-Ekiti, Nigeria*
[2]*Department of Acturial Science. Ekiti State University, Ado-Ekiti, Nigeria*

Submitted: 10-07-2021 Revised: 23-07-2021 Accepted: 26-07-2021

ABSTRACT:Survival time analysis is the analysis of time to event. It is particularly used to identify risks involved in the survival data. A special type of data has to be analyzed with special methods. When survival times are analyzed without the use of special techniques, and if underlying assumptions were not taken into an account, then faulty interpretation would result. Tuberculosis (TB) is an infectious disease which has been persistent with humans throughout known history. The aim of this study is to establish the influence of socioeconomic, demographic as well as the environmental variables on tuberculosis patients in Nigeria. Non parametric (Kaplan Meire) and semi parametric (Cox proportional hazard) methods in Survival analysis were used in analyzing the data**.**Hazard ratio of different variables was determined using cox proportional hazard regression.Resultobtained by the PHreg procedure indicated tuberculosis category gender and age of patients as factors influencing tuberculosis patient. In this study we concluded that the accepted year - to-treatment hazard ratio for distinct concentrations in relation to the reference risk (year 2018) is 1.147, 0.669, 0.912 for years 2015, 2016, 2017 respectively. This means that patients admitted in 2018 have a brief survival time (i.e. will die more quickly from TB) relative to the baseline hazard (year 2018), and patients admitted for therapy in 2016 and 2017 also have a lower risk and survive longer than the baseline hazard. It has been noted, however, that year does not make a significant contribution to the model.

Keywords Tuberculosis, Survival function, Hazard ratio, Cox regression model

I. INTRODUCTION

Tuberculosis (TB) is an infectious disease which has been persistent with humans throughout known history [1]. TB is a disease caused by mycobacterium that mainly affects the lungs spreading to other organs such as brain, skin and other viscera. The causative organisms as outlined by Assam [2] are Mycobacteriumtuberculosis and Mycobacterium bovis. Tuberculosis (TB) remains a major global public health problem and is second to human immunodeficiency virus (HIV) as the leading cause of death due to a single infectious agent in the world and loss of healthy life years in the productive age group [3, 4]. TB affects all the predicted field of quality of life such as general health perceptions, corporal sense, psychological health, mental peace and functionality of physical and social roles [5].

In 2009, U.S. Global Health policy in Global Tuberculosis Epidemic declared that about one-third of the world's population or two billion people carried the TB bacteria, and more than 9 million of who become sick each year with 'active' TB which could be spread to others [6]. In low and middle- income countries (LMICs), TB stands third among the leading cause of adult mortality after HIV and ischemic heart disease [7]. Despite the down trend in the incidence and prevalence of TB, every continent still report new cases especially Africa and south-east Asia [8]. The highest rates of TB cases are found in countries where poverty, crowding and insufficient health care programs are common problems [9]. Nigeria is one of the countries in sub-Saharan Africa, noted to be saddled with a high prevalence of the disease [10]. In year 2007, Nigeria ranked fourth in the world and first in Africa with respect to the WHO estimated number of TB cases. Unfortunately, a 2008 report estimated total TB cases in Nigeria as 922,575, and was ranked 3rd (behind India and China) on the list of high-burden countries [11]. Furthermore, as at 2007, WHO estimated that Nigeria had 460,000 cases of all form of TB, a TB prevalence of 521/100,000 population, 195,000 new smear positive cases, incidence rate (all cases) of 311/100,000 per year, and the occurrence rate (new smear positive) of 131/100,000 per year. Further estimates include the prevalence of all forms of TB in HIV of 42/100,000, and a death rate of 93/100,000 population per year (138,000 deaths/year) (WHO, 2009).

Without the implementation of proper control measures, WHO estimates that between 2000 and 2020, nearly one billion people will be newly infected, 200 million people will get sick and 35 million will

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

die from TB (WHO, 2008). Health seeking behavior and the perceived knowledge on cause of TB among community member is very critical and may reduce or increase the transmission of the disease. Certain local practices, beliefs, such as illness representation of the illness character and shame related to it, and failure to recognize symptoms early may delay diagnosis hence increasing the spread of the disease in the community [12]. Therefore, public awareness is essential for the reduction of both mortality and morbidity of TB [13]. It is well established that good public awareness correlates well with the early detection of disease. Knowledge plays a vital role in influencing the behavior and practices of the individuals. Early detection and diagnosis of TB can cause a decrease in TB mortalities and occurrence [14, 15]. The need for population based studies in order to design appropriate tuberculosis education, should not therefore be understated if the global targets for case detection and treatment outcome are to be achieved [16]. This study used data from the University Teaching Hospital (UITH), Ilorin Nigeria. Non parametric and semi parametric methods were used in analyzing the data**.**

II. MATERIALS AND METHODS

The hazard feature h(t), given no prior occurrences, is the instantaneous speed at which incidents happen.

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} \quad (1)$$

Where T is the survival time random variable, S(t), Survival function, and Δt is the change in specific value of time t

2.1 Log Rank Test

The Log rank test also known as Mantel-Haenszel test is a large sample chi-square test that uses a statistic as its test criterion that compares the overall Kaplan-Meier (KM) curves. It applies to information where there is progressive censorship and early and late errors are given equal weight. Like many other statistics, this statistics uses the observed and expected cell count over outcome categories where each of the ordered failure times for the entire set of data being analyzed defines the categories for the log rank statistics. It is also a test used to compare distribution of survival for two or more groups and assumes parallel functions of hazard for the two groups. When comparing two groups, a 1-degree-free statistics (known as Mantel-Haenszel statistic or log rank test statistics) is formed using the sum of observed minus expected counts for one of the two groups over all failure times. It can also be calculated by separating the square of the two observed sum less anticipated score by variance of the measured sum less anticipated score for one group. The test statistic can be expressed as:

$$\frac{(O_i - E_i)^2}{\mathrm{Var}(O_i - E_i)}, i = 1,2. \quad (2)$$

$O_i - E_i$ = observed cell counts score minus expected cell counts score

Var $(O_i - E_i)$ = variance of observed cell counts score minus expected cell counts score

The log rank test, which is a family of the non-generalized Wilcoxon test, is not very efficient when the two distributions vary but their risk functions or survival functions are crossed, none of them is strong to identify any differences and it will be reasonable to consider other tests. It can also be expanded to compare two or more survival groups ' failure times.

2.2 Cox regression model

The model of Cox regression is one of the most used modelling techniques in survival analysis because it is robust; a close approximation of the outcomes can be made for the right parametric model based on the outcomes of using the model of Cox. Cox model also provides a non-negative hazard rate at all times. The exponential portion of the hazard function is attractive as it always causes the fitted model to offer non-negative estimated risks. Another attractive feature of the Cox model is that it is still feasible to predict the β's in the exponential portion of the model despite not specifying the model's baseline risk portion. The hazard ratio, which is an impact metric, is calculated without the baseline hazard function needing to be bothered.

The general form of the Cox model is given by;

$$h(T, x) = h_0(t)\exp\left(\sum_{i=1}^{k} \beta_i X_i\right) \quad (3)$$

Where $h_0(t)$ is the baseline hazard and the β_i's are the regression coefficient with i = 1,2,3, … , p and X's are the explanatory variables.

III. SUMMARY OF DATA PRESENTATION

The information gathered for this work is a secondary data acquired from the records of 385 TB patients registered at University of Ilorin Teaching Hospital (UITH), Ilorin Nigeria. The data gathered are based on duration of stay, age, gender, types of TB (Pulmonary, additional pulmonary) for a period of four (4) years, i.e 2015-2018.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

Table 1.0: Frequency table showing gender of patients

GENDER	Frequency	Events	Censored		Valid Percentage	Cumulative Percentage
			N	Percentage		
Male	196	71	125	63.8%	50.9%	50.9%
Female	189	55	134	70.9%	49.1%	100.0%
Total	385	126	259	67.3%	100.0%	

Table 2.0: Frequency table showing Tb category of patients

TB CATEGORY	Frequency	Events	Censored		Valid Percentage	Cumulative Percentage
			N	Percentage		
Pulmonary	336	105	231	68.8%	87.3%	87.3%
Extra – pulmonary	49	21	28	57.1%	12.7%	100.0%
Total	385	126	259	67.3%	100.0%	

Table 3.0: Frequency table showing age groups of Tb patients

AGE	Frequency	Events	Censored		Valid Percentage	Cumulative Percentage
			N	Percentage		
<15 years	33	4	29	87.9%	8.6%	8.6%
16-24 years	31	6	25	80.6%	8.1%	16.6%
25-34 years	110	30	80	72.7%	28.6%	45.2%
>35 years	211	86	125	59.2%	54.8%	100.0%
Total	385	126	259	67.3%	100.0%	

Table 4.0: Frequency table showing status of Tb patients

STATUS	Frequency	Events	Censored		Valid Percentage	Cumulative Percentage
			N	Percentage		
Alive	219	198	21	9.6%	56.9%	56.9%
Dead	126	105	21	16.7%	32.7%	89.6%
Dama	37	30	7	18.9%	9.6%	99.2%
Trans	3	3	0	0.0%	0.8%	100.0%
Total	385	336	49	12.7%	100.0%	

Table 5.0: Frequency table showing year of admission

YEAR	Frequency	Events	Censored		Valid Percentage	Cumulative Percentage
			N	Percentage		
2015	112	44	68	60.7%	29.1	92.5
2016	133	36	97	72.9%	34.5	34.5

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

2017	111	37	74	66.7%	28.8	63.4
2018	29	9	20	69.0%	7.5	100.0
Total	385	126	259	67.3%	100.0%	

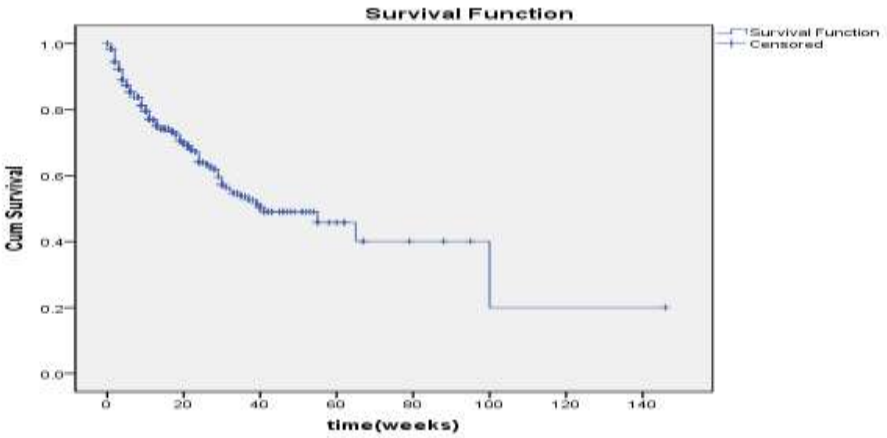

Fig 1: The survival curve for Tuberculosis patients

Figure 1 shows the survivorship function of tuberculosis patients with their respective number at risks at different time interval.

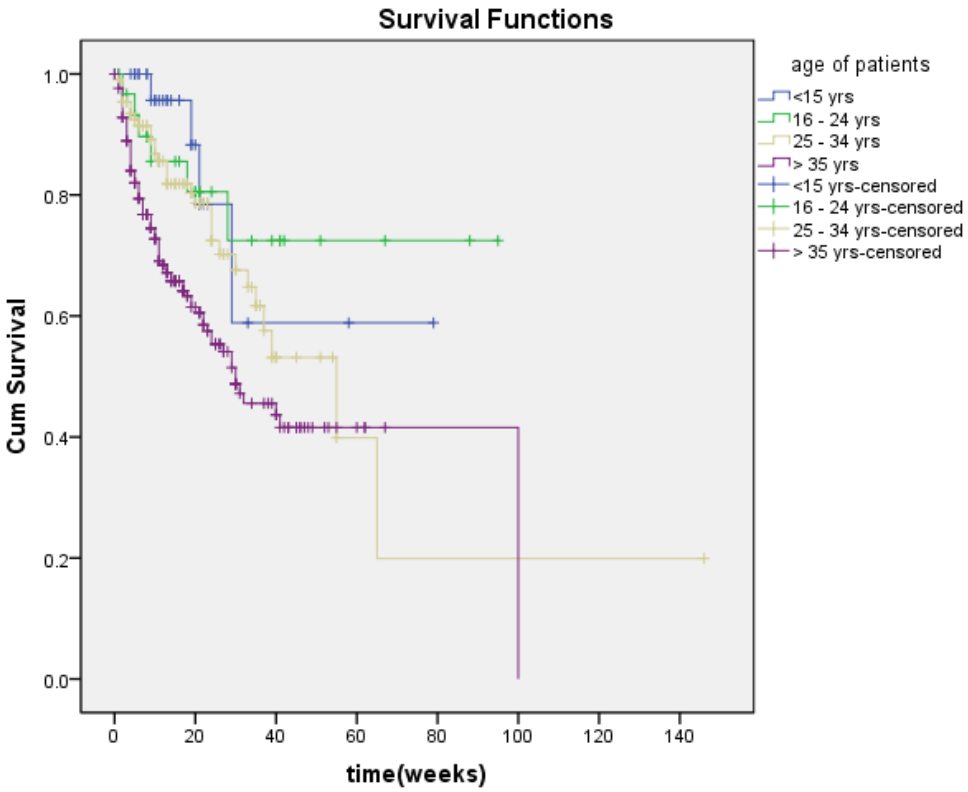

Fig 2: The survival curve based on age of Tuberculosis patients

Figure 2 shows the survivorship function based on the age distribution of tuberculosis patients with the number at risk at several time points.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

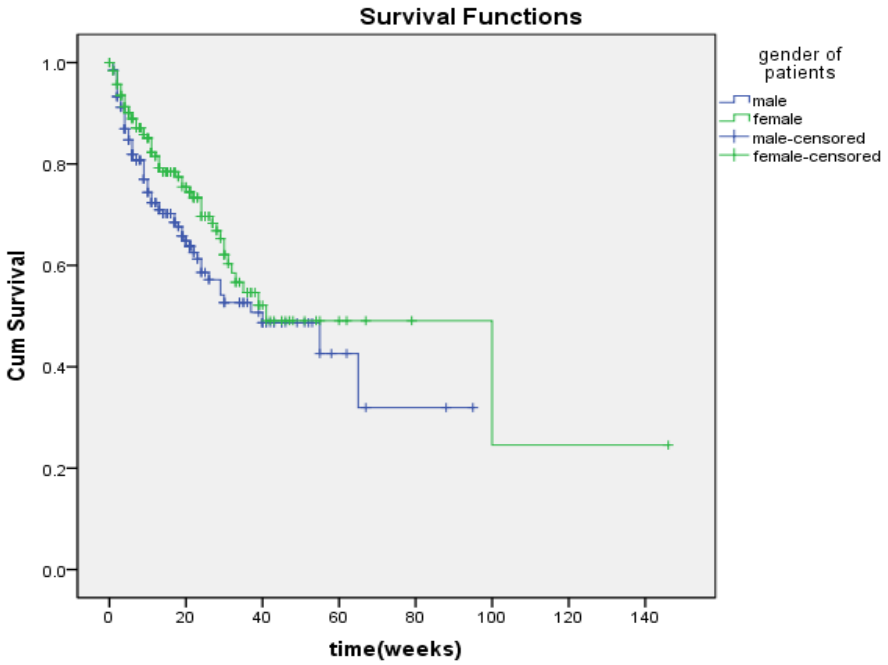

Fig 3: The survival curve based on gender of Tuberculosis patients

Figure 3 reveals the survivorship function on the gender of tuberculosis patients which include the number at risk at several time points.

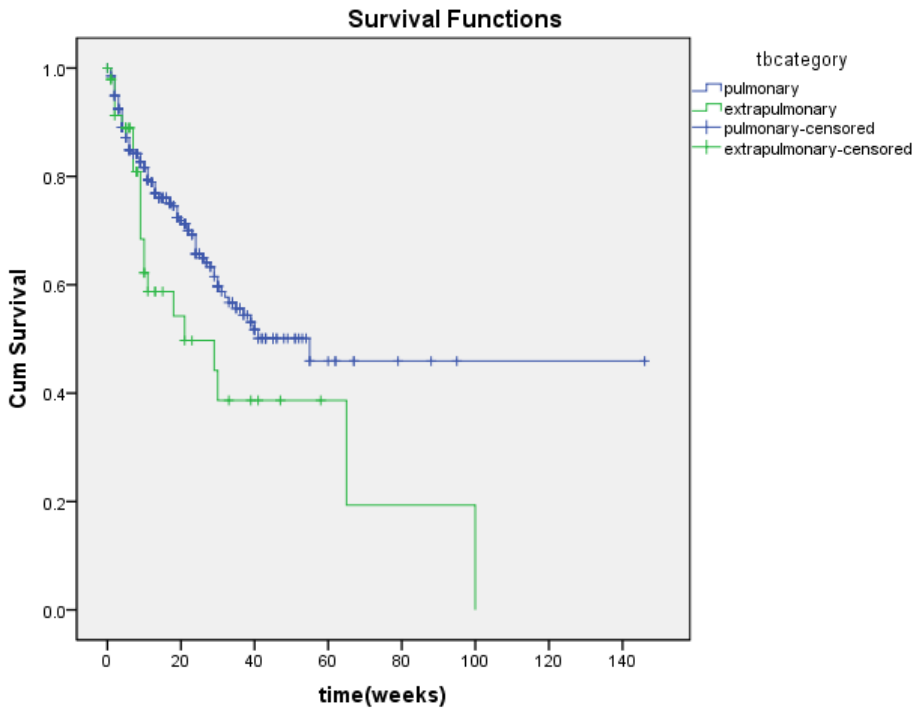

Fig 4: The survival curve based on type of Tuberculosis disease

Figure 4 indicates the survivorship function on types of tuberculosis with their respective number at risk in different time points.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

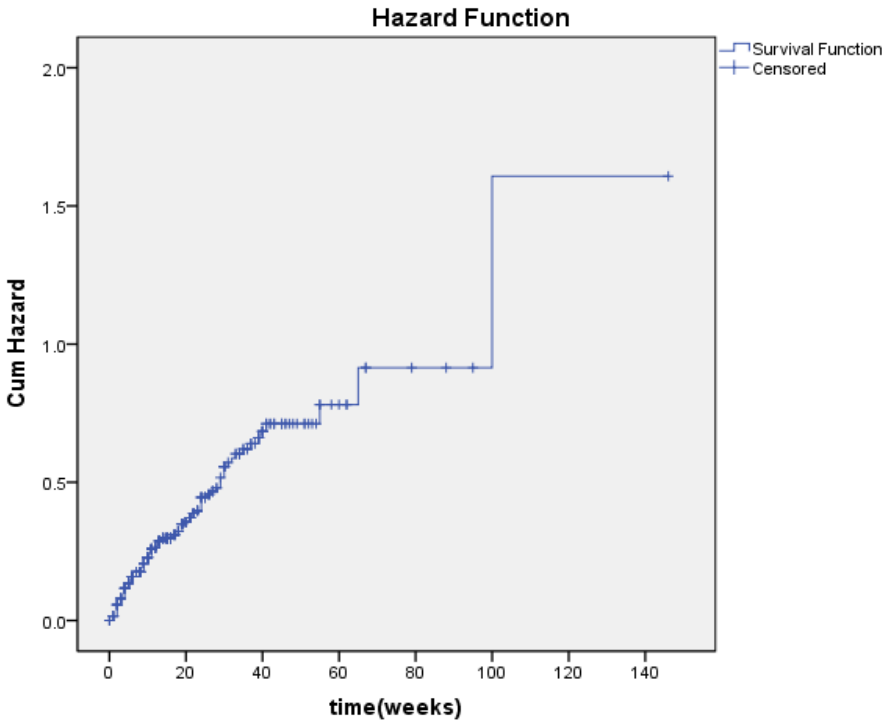

Fig 5: The hazard function curve

Figure 5 indicates the hazard function of tuberculosis patients with their respective number at risk in different time points.

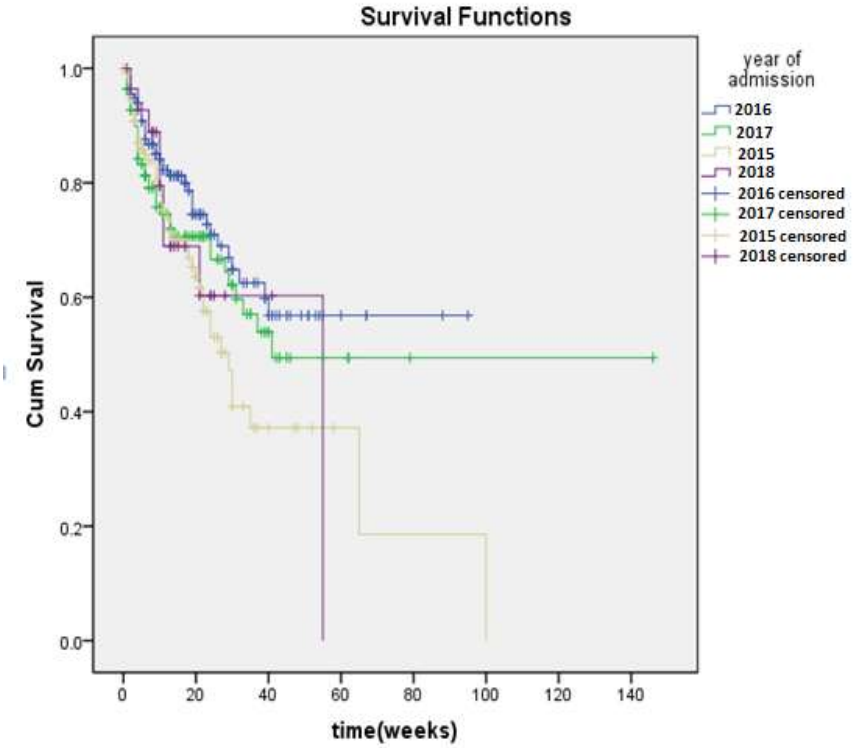

Fig 6: Survival curve based on year admitted

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

Figure 6 reveals the survival function on the year of admission of tuberculosis patients with their respective number at risks atdifferent time interval.

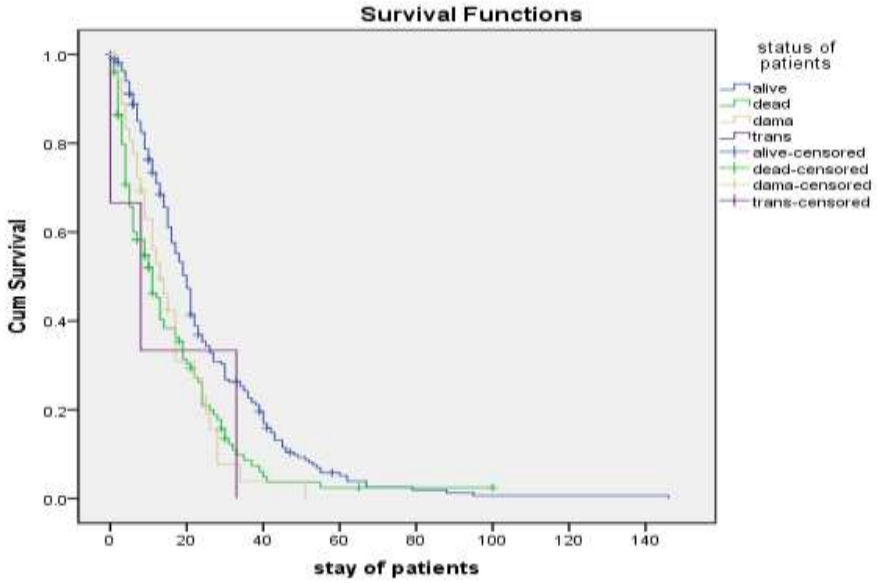

Fig 7: Survival curve based on status of Tuberculosis patients

Figure 7 reveals the survivorship function on the status of tuberculosis patients with their respective number of event (dead) gotten at different time interval.

3.1. KAPLAN-MEIER ESTIMATES

Table6.0: Summary Statistics of Survival Time Using KM- Method

Records	n.max	n.start	Events	Median
385	385	385	126	41

Table 7.0: Kaplan-Meier Estimates of Survival Time

time	n.risk	n.event	survival	std.err	Lower 95%CI	upper95% CI
1	383	6	0.984	0.00635	0.9720	0.997
2	372	15	0.945	0.01174	0.9219	0.968
3	352	8	0.923	0.01371	0.8967	0.950
4	338	12	0.890	0.01616	0.8593	0.923
5	319	6	0.874	0.01724	0.8405	0.908
6	303	7	0.853	0.01846	0.8181	0.890
7	287	5	0.839	0.01929	0.8016	0.877
9	261	8	0.813	0.02073	0.7733	0.855
10	243	5	0.796	0.02161	0.7549	0.840
11	231	7	0.772	0.02280	0.7286	0.818
12	215	1	0.768	0.02298	0.7247	0.815
13	208	5	0.750	0.02386	0.7046	0.798
14	195	2	0.742	0.02423	0.6963	0.791
17	168	2	0.733	0.02473	0.6865	0.784
18	158	2	0.724	0.002528	0.6763	0.775
19	152	4	0.705	0.02635	0.6553	0.759
20	142	1	0.700	0.02663	0.6499	0.754
21	135	2	0.690	0.02722	0.6384	0.745

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

22	120	2	0.678	0.02796	0.6256	0.735
23	112	1	0.672	0.02836	0.6189	0.730
24	106	5	0.641	0.03036	0.5837	0.703
26	94	1	0.634	0.03079	0.5761	0.697
27	89	1	0.627	0.03126	0.5682	0.691
28	84	1	0.619	0.03176	0.5599	0.685
29	81	3	0.596	0.03323	0.5345	0.665
30	77	3	0.573	0.03454	0.5091	0.645
31	67	1	0.564	0.03506	0.4997	0.637
32	65	1	0.556	0.03558	0.4902	0.630
33	64	1	0.547	0.03607	0.4807	0.623
35	58	1	0.538	0.03666	0.4704	0.614
37	52	1	0.527	0.03738	0.4589	0.606
39	48	1	0.516	0.03819	0.4466	0.597
40	43	1	0.504	0.03914	0.4331	0.587
41	37	1	0.491	0.04039	0.4176	0.577
55	15	1	0.458	0.04919	0.3710	0.565
65	8	1	0.401	0.06870	0.2863	0.561
100	2	1	0.200	0.14577	0.0481	0.834

3.2 Log Rank Test

3.2.1 Comparison Of Survival Experience Based On Gender.

Table 8. Log-Rank Test Comparing Gender

Gender	N	Observed	Expected
Male	196	71	61.5
Female	189	55	64.5

$\chi^2 = 2.9$ p-value= 0.0859

The survival functions for the gender are not different from each other. In other words, they have similar survival functions.

Table 9. Log-Rank Test Comparing Age Of Patients

Patient's Age	N	Observed	Expected
<15 years	33	4	9.75
16-24 years	31	6	12.12
25-34 years	110	30	39.21
>35 years	211	86	64.92

$\chi^2 = 15.9$ p-value= 0.00118

In the table 9 above, it is seen that the survival functions for the different age groups are not similar.

Table 10. Log-Rank Test Comparing Category Of Tuberculosis

Tb category	N	Observed	Expected
Pulmonary	3336	105	112.2
Extra-pulmonary	49	21	13.8

$\chi^2 = 4.4$ p-value= 0.0359

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 2981-2990 www.ijaem.net ISSN: 2395-5252

In the table 10, the Survival functions of the different Tb categories are not similar.

Table 11: Log-Rank Test Comparing Year Admitted

Year	N	Observed	Expected
2015	112	44	33.28
2016	133	36	47.42
2017	111	37	37.17
2018	29	9	8.12

$\chi^2 = 6.5$ p-value= 0.0897

According to table 11, the survival functions for the different years are similar.

Table 12.Cox Proportional Hazards Model Results

Covariates	Coef	exp(coef)	se(coef)	z	P
PATIENT'S AGE <15 years	--1.6958	0.183	0.534	-3.176	0.0015
PATIENT'S AGE 16-24 years	--1.0066	0.365	0.432	-2.328	0.0200
PATIENT'S AGE 25-34 years	--0.5728	0.564	0.213	-2.688	0.0072
Gender	0.4614	1.586	0.187	2.461	0.0140
Tb category	--0.6821	0.506	0.276	-2.468	0.0140
Year 2008	0.1370	1.147	0.375	0.365	0.7100
Year 2009	--0.4014	0.669	0.395	-1.016	0.3100
Year 2010	--0.0926	0.912	0.389	-0.238	0.8100

The hazard ratio for Gender adjusted for other covariates, compare to female is 1.586 as shown in (Table 12.0), since 1 is less than this ratio. Therefore, male patients were suffering from TB as compared to female patients. To ascertain if covariates contributed significantly to the model, their p- values were compared to the significance level (0.05) and it was observed that the risk of dying based on gender is significant since the p-value (0.0140) is less than 0.05.

IV. DISCUSSION

A number of techniques have been used to describe the 385 subjects ' survival distributions. The following observations were made and interpreted from the study; it was concluded that patients with Pulmonary Tb survive longer with a median survival time of 55 weeks than patients with additional pulmonary Tb with a median survival time of 21 weeks.

The Survival curves show distinct age group survival experiences. That is, the survival time of patients < 15 years of age differed from those 16-24, 25-34, and > 35 years of era. The log-rank test was also performed to compare the survival function of the different co-variate concentrations. It was noted that the survival experiences of the distinct gender and distinct years admitted for therapy at 0.05 point of importance were not substantially distinct. However, the survival experiences of the distinct classifications of the era of the patient and the status of Tb differed considerably from each other at 0.05 point of meaning, by fitting cox regression model, the estimated coefficients of regression (coef) and the risk ratio (exp (coef)) were acquired between the two covariate groups. Meanwhile, based on its proximity to 1, it is thus interpreted. The Gender-adjusted hazard ratio for other covariates relative to female is 1,586 as obtained in (Table 12.0), since 1 is less than this ratio. Therefore, male patients were suffering from TB as compared to female patients and the risk of male patients getting the event from TB is 1,586 times that of female patients.

Consequently, the risk ratio for the distinct covariate age concentrations of the patient relative to the reference hazard (patient age > 35 years) is 0.183, 0.365, 0.564 for ages < 15, 16-24 years, 25-34 years respectively. This means that patients in these age groups have a lower risk of TB dying and a longer survival time than patients who are > 35 years of age. It has also been noted that the age groups make a significant contribution to the model. The risk of dying from pulmonary TB is 0.506 times relative to extra pulmonary TB in the same vein; this means that

patients with pulmonary TB have a lower risk of dying from Tb with a lengthy survival time. This also makes a significant contribution to the model.

V. CONCLUSIONS

In conclusion, the accepted year - to-treatment hazard ratio for distinct concentrations in relation to the reference risk (year 2018) is 1.147, 0.669, 0.912 for years 2015, 2016, 2017 respectively. This means that patients admitted in 2018 have a brief survival time (i.e. will die more quickly from TB) relative to the baseline hazard (year 2018), and patients admitted for therapy in 2016 and 2017 also have a lower risk and survive longer than the baseline hazard. It has been noted, however, that year does not make a significant contribution to the model.

Acknowledgements

We are very grateful to experts for their appropriate and constructive suggestions to improve this template.

REFERENCES

[1] Daniel T.M (2006) The History of Tuberculosis. Respiratory Medicine, 100(11):1862-1870

[2] Assam JPA, Beng VP, Cho-Ngwa F et al. Mycobacterium tuberculosis is the causative agent of tuberculosis in the southern ecological zones of Cameroon, as shown by genetic analysis BMC Infectious Diseases 2013,13:43.

[3] Vidhani M and Vadgama P. Awareness regarding pulmonary tuberculosisa study among patient taking treatment of tuberculosis in rural Surat, Gujarat. Natl J Med Res 2012; 2: 452-455

[4] Baveja, SM., Dalal, PJ. The attitude to tuberculosis patients among medical undergraduates and Awareness of the national tuberculosis control programme .J Acad Med Sci 2012;2:68-72.

[5] Unalan, D., Ferhan, S., Osman, .C., Mustafa,. B., and Ahmet, O. Does the quality of life different in patients with active and Inactive tuberculosis. Indian Journal of Tuberculosis 2008 55(3):127-37.

[6] WHO: Global Tuberculosis Control: Epidemiology, Strategy, Financing, WHO Report 2009, WHO/HTM /TB2009. 411: 1-309.

[7] Sreeramareddy, CT., Harshakumar, HN., Arokiasamy, JT. The Prevalence of self-reported tuberculosis, knowledge about tuberculosis transmission and its control among adults in India: results from a nation-wide crosssectional household survey. BMC Infect Dis 2013;13:16.

[8] WHO: Tuberculosis; Global Tuberculosis control report 2008. Fact sheet N 104

[9] Farmer, TA. Prospects for reducing tuberculosis incidence and mortality, 2005, 72(2): 200-21.

[10] Egah, TE. and Okoli, CU., Tuberculosis in Jos Nigeria: Laboratory Report at Jos University Teaching Hospital (JUTH), A 9 year Review of. Nigeria Med Pract, 46:36-46, 2004.

[11] Kaiser Family Foundation. Global Health Facts: Adult HIV/AIDS Risk Rate (Aged 15-49). Available at: http://www global health facts org/topic jsp?i=3. Accessed April 14, 2010.

[12] Woith, WM., Larson, JL. Delay in seeking treatment and adherence to tuberculosis medications in Russia: A survey of patients from two clinics. International Journal of Nursing Studies 2008;45(8):1163-1174

[13] Gijs, E., Mario, CR., and Dermot, M. Scale Up: Reaching Targets in Global Tuberculosis Control. The Lancet 2004, 363, 814-819

[14] Suárez, PG., Watt, CJ., Alarcón, E., Portocarrero, J., Zavala, D., Canales, R., Luelmo, F., Espinal, MA. and Dye, C. The Activeness of Tuberculosis in Response to 10 Years of Intensive Control Effort in Peru. The Journal of Infectious Diseases 2001, 184, 473-478.

[15] Nathan, S. K., Thomas, G. H., & McConnick, D. S. Sources, diagnosis and prevention of TB in the tropics. Journal of Medical Association, 23, 33-39, 2003.

[16] Kochi, A. The global tuberculosis situation and the current control strategy of the World Health Organization. Tubercle 1991;72(1):1-6.