RESEARCH ARTICLE

JENER Journal of Empirical and Non-Empirical Research
Volume 2, Issue 1, Page 312-318 | Article No. 36
DOI: 10.4898/jener.v2i1.a36
ISSN: 7895-9216
Journal homepage: https://jenerjournal.com

JENER Journal
JENER Journal of Empirical and Non-Empirical Research

On The Survival of Tuberculosis Patients in Makurdi, Benue State Using Kaplan Meier Estimator: A Case Study of Madonna Hospital Makurdi

M. K. Adamu [1]*, S. S. Afiagh[2]

[1]*Department of Statistics, Joseph Sarwuan Tarka University, Makurdi*
[2]*Statistical Computing Laboratory, Department of Statistics, Jospeh Sarwuan Tarka University, Makurdi*
*Corresponding Author: M. K. Adamu, Email: adamu.michael@uam.edu.ng

APA Citation and Referencing: Adamu, M.K., & Afiagh, S.S. (2026). On The Survival of Tuberculosis Patients in Makurdi, Benue State Using Kaplan Meier Estimator: A Case Study of Madonna Hospital Makurdi. *JENER Journal of Empirical and Non-Empirical Research*, 2(1), 312-318

ARTICLE INFORMATION

Article history:
Published on 23rd Jan 2026

Keywords:
Survival analysis
Kaplan Meier Estimator
Tuberculosis

ABSTRACT

This research work was carried out to model the survival of Tuberculosis patients in Makurdi, Benue State. Kaplan-Meier estimator was used to estimate the median of the distribution of Tuberculosis patients' survival times following their recruitment into the study. Age, sex, HIV status, and results of the treatment of 98 Tuberculosis patients were the variables used in the research. The result showed that 51.02% of the population were male and 48.98% were female. The results also showed that 30.6% of the Tuberculosis patients were HIV negative, 66.3% were HIV positive and 30.1% of HIV status were unknown. The mean age of the tuberculosis patients is 42 years. The overall median survival time for Tuberculosis patients was 27 days. This indicates that 50% of the Tuberculosis patients survived longer than 27 days after being diagnosed with the disease. The log-rank test showed that the survival time between males and females was not significantly different at $\propto = 0.05$, and it was concluded that since most of the population survive pass 27 days, the treatment regimen is a good one.

1. Introduction

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. Time may be years, months, weeks, or days from the beginning of follow-up, or the age of an individual when an event occurs. An event may be death, disease incidence, relapse from remission, recovery, or any designated experience of interest. Although more than one event may be considered, survival analysis usually assumes only one event of designated interest; when multiple events are involved, the problem may be a recurrent event or a competing risk problem. The time variable is called survival time, and the event is typically referred to as a failure, even though failure may represent a positive event such as return to work. Examples of survival analysis include studies of leukaemia patients in remission, disease-free individuals developing heart disease, elderly populations followed for survival, recidivism studies of parolees, and survival after heart transplantation (Kleinbaum et al., 2007). A unique feature of survival data is censoring, where not all patients experience the event by the end of the observation period, so actual survival times for some patients are unknown and must be accounted for to allow for valid inferences. Survival times are usually skewed, limiting methods that assume a normal distribution. Appropriate statistical methods include nonparametric and semiparametric methods, specifically the Kaplan–Meier estimator, log-rank test, and Cox proportional hazards model, which are the most commonly used techniques in medical literature (Schober et al., 2018). Survival analysis has been used to solve problem from a record of 337 confirmed cases of tuberculosis patients in Monrovia, the capital of Liberia, 2015. Kaplan-Meier analysis and the log-rank test were used to assess the differences in survival among the patients, while Cox regression model was used for multivariate analysis. The mean period of the follow-up of patients was 10 months. In the 337 patients, 33 (9.8%) died, the 21-month survival rate was 90.2%. The results of multivariate Cox regression analysis show that overcrowding, former smoking, current smoking, multidrug-resistance tuberculosis were risk factors for death during anti-tuberculosis treatment. The results of binary logistic regression analysis show that extra-pulmonary, family history of TB and current smoking were risk factors for multidrug-resistant tuberculosis. In conclusion, overcrowding, smoking, MDR-TB were important risk factors and negatively affected the survival rates of TB patients in Liberia (Carter et al., 2021). Most survival analyses must consider a key analytical problem called censoring. Censoring occurs when we have some information about individual survival time, but we don't know the survival time exactly. Right-censored is when true survival time is equal to or greater than observed survival time. Left censored is when true survival time is less than or equal to the observed survival time. Interval-censored is when true survival time is within a known time interval (Kleinbaum et al., 2007). Tuberculosis (TB) is a chronic respiratory infectious disease caused by Mycobacterium tuberculosis and remains one of the top 10 causes of death worldwide. China has approximately 900,000 new patients with TB annually (Liao et al., 2017). In 2018, 87% of new TB cases

occurred in the 30 high TB burdened countries, and Liberia is among the thirty high TB burdened countries with an estimated incidence rate of 308 per 100,000 population (Tekie, 2018). The World Health Organization's 2019 Global Tuberculosis Report indicated 10 million new patients with TB worldwide and 1.5 million deaths in 2018 (WHO, 2019). The rise in cases of tuberculosis in Makurdi has posed a serious health emergency, and the need to carefully study and curtail the magnitude of spread of Tuberculosis becomes necessary as a major step for positive intervention strategy. The aim of this work is to investigate the risk factors on the reported cases of Tuberculosis infection in Makurdi, using time-to-event analysis Kaplan Meier estimator, to estimate the survival time including median time of survival, compare the patterns of the event times of tuberculosis in the patients, and compare the event survival time for male and female and the relationship between gender of the event and HIV status of the event. This research work will help the general public on how to relate with the affected, health workers, and the government intervention strategy, and also to study possible cure recommendation for Tuberculosis in order to prevent the transmission and spread of the disease.

2. Literature Review

Ajagbe et al. (2014) carried out a retrospective cohort research study which tries to analyse the survival of adult TB disease of patients that was treated in County Cork, Ireland. We included only patients that were 18 years and above as at the time of diagnosis and had diagnosis and treatment between January 2008 – December 2012. A total of 647 patients were included in this study. There was no significant difference in survival curves of male and female and among age group. Women had a longer ST compared to men, but men had a higher hazard rate compared to women. Akessa et al. (2015) carried out a study on survival analysis of loss to follow-up treatment among Tuberculosis patients at Jimma University Specialized Hospital, Jimma, Southwest Ethiopia. A total of 510 patients who initiated TB treatment from February 2011 to September 2013 were included in the study. Sixty-nine patients (13.5%) were lost to follow-up of their treatment. The probability of survival after TB treatment at the end of follow-up was 85.48%. Yigzaw and Demeke (2018) carried out a retrospective study in seven hospitals having MDR-TB treatment center around Amhara region of Ethiopia, from September 2015–February 2018. The recovery time of patients around Amhara region was 21 months. Male TB patients, patients with co-morbidity, smoking and clinical complication were experiencing longer recovery time than that of the counter groups. Mulugeta (2018) carried out a retrospective cohort study titled modeling the survival of TB patients in Eastern region of Tigray in Adigrat and Wukrogeneral hospitals in Ethiopia. On the 397 patients studied over the specified period, 23 (5.8%) had died. Significant survival difference was observed among gender, residence, HIV status, treatment category and age category of patients. Balaky et al. (2019) carried out a research work titled "Survival analysis of patients with tuberculosis in Erbil, Iraqi Kurdistan region". Of 728 patients with TB, 50 (6.8%) had died. The 12-month survival rate of their study was 93.1%. A statistically significant difference was noticed in the survival curves of different age groups. Abbara et al. (2018) carried out a retrospective cohort study comparing the clinical, radiological and demographic characteristics of TB patients aged ≥65 years with TB patients aged 18–64 years in the UK. Older patients aged ≥65 years with TB had fewer "classical" clinical and radiological presentations of TB, which may explain longer times to starting treatment from symptom onset compared with younger patients aged <65 years. Xie et al. (2020) prospectively analysed the effect of supervision and treatment on newly diagnosed TB patients in Haihe Hospital, Tianjin from 2014–2017. A total of 7,032 patients received standardized antituberculosis treatment. A total of 321 (4.56%) patients died and 6,711 (95.44%) survived; 50.47% of the patients died in the intensive period 2 months treatment. Bolarinwa and Michael (2020) carried out a study aimed at fitting Cox-proportional hazards model to Tuberculosis data from the Federal Medical Centre, Bida, Nigeria. Model (G+C), with gender and occupation as covariates produced the least AIC of 618.597. Gender and occupation constituted the best subset of covariates that explained survival of TB patients. Liu et al. (2021) investigated the factors associated with PTB death in Zhejiang Province, China. The 1-year, 3-year, and 5-year survival percentages were 46.6%, 25.1%, and 14.5%, respectively. Considering that 86.6% of participants were older than 60 years, it is suggested that weakened individual immunity and some existing comorbidities may aggravate the PTB infection, onset, and even death.

3. Methodology

Source of Data

The data for this study was a secondary data obtained from Madonna hospital, Wurukum, Makurdi, Benue State. It consisted of 98 patients diagnosed of Tuberculosis disease and received treatment from 2018 to 2022 with a treatment period of six (6) months. SPSS (Statistical Package for the social sciences) was the software used to analyze the data.

Survival Analysis

Survival Function

Given that T is the survival time and t is the specific value of interest for the random variable T;
The survival function is denoted by $S(t)$.

$$S(t) = \Pr(T > t) = 1 - F(t)$$

where F(t) is the probability density function.
The survival function gives the probability that a subject will survive past time t.

The Hazard Function

$h(t)$ is the instantaneous rate at which events occur, given no previous events.

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

• The cumulative hazard describes the accumulated risk up to time t,

$$H(t) = \int_0^t h(\mu)\, du$$

Chi-square

Pearson's chi-square test uses the formula below to calculate the test statistic.
chi-square written as (χ^2):

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where:

- χ^2 is the chi-square test statistic
- Σ is the summation operator (it means "take the sum of")
- is the observed frequency
- E is the expected frequency

Odds Ratio:

The odds ratio tells us the ratio of the odds of an event occurring in a treatment group to the odds of an event occurring in a control group. It is calculated as:

$$\text{OR} = \frac{\text{odds that a case was exposed (A/C)}}{\text{odds that a control was exposed(B/D)}}$$

$$\text{OR} = \frac{(\text{AD})}{(\text{BC})}$$

Logrank Test

Definition: $T_1 < T_2 < \cdots < T_k$ are distinct failure times
$Y_i(T_j) =$ Number of persons in group i at risk at T_j
$Y(T_j) = Y_0(T_j) + Y_1(T_j)$, the total number of subjects at risk T_j
$d_{ij} =$ Number of failures in group i at T_j
$d_j = d_{0j} + d_{1j}$ total number of failures at T_j
The information at time T_j can be summarized in the following 2x2 table:

Table 1: The Log Rank Test Statistics.

	Observed to fail at T_j		At risk at T_j
Group 0	d_{0j}	$Y_0((T_j) - d_{0j}$	$Y_0(T_j)$
Group 1	d_{1j}	$Y_1((T_j) - d_{1j}$	$Y_1(T_j)$
	d_j	$Y(T_j) - d_j$	$Y(T_j)$

The log-rank Test statistic is given as;

$$Z = \frac{\sum_{j=1}^{k}(O_j - E_j)}{\sqrt{\sum_{j=1}^{k} V_j}} \sim N(0,1) under\ H_0$$

Where;
$O_j = d_{1j}$: Observed number of failures

$E_j = d_j \frac{Y_1(T_j)}{Y(T_j)}$: expected number of failures

$V_j = \frac{Y_0(T_j)Y_1(T_j)d_j(Y(T_j)-d_j}{Y(T_j)^2\,(Y(T_j)-1)}$: Variance of the observed number of failures.

Kaplan Meier Estimator

Kaplan Meier estimator is given by;

$$\hat{S}(t) = \prod_{t_i \le 1}\left(1 - \frac{d_i}{n_i}\right)$$

Or

$$\prod_{t_i \le 1} \frac{n_i - d_i}{n_i}$$

With a variance given by;

$$\hat{S}^2(t) = \sum_{t_i \le 1} \frac{d_i}{n_i(n_i - d_i)}$$

Where;
$\frac{d_i}{n_i}$ is the proportion or estimated probability of having an event on the interval from t_{i-1} to t_i
$1 - \frac{d_i}{n_i}$ is the proportion or estimated probability of surviving that same interval
t_i is the duration of study at point i
d_i is the number of deaths up to point i
n_i is the number of individuals at risk prior to t_i

Assumptions of Kaplan Meier Estimator
- The survival probability is the same for censored and uncensored subjects.
- The likelihood of the occurrence of the event is the same for the participants enrolled early and late.
- The event is assumed to occur at the defined time.

Test of Hypotheses
We compared the event time for male and female by testing the Hypothesis as follows;
H_0: There is no difference between male survival time and female survival time.
H_1: There is a difference between male and female survival time.
Our $\propto = 0.05$
Reject H_0 if P-value ≤ 0.05 and accept H_1 which means there is significant difference.
Reject H_1 if P-value > 0.05 and accept H_0 which means there is no significant difference.
We also compared if there is a relationship between gender of the events and HIV Status of the events by testing the Hypothes is as follows;
H_0: There is no relationship between gender and HIV Status of the events.
H_1: There is a relationship between gender and HIV Status of the events.
Our $\propto = 0.05$
Reject H_0 if P-value ≤ 0.05 and accept H_1 which means there is relationship between gender and HIV Status of the events.
Reject H_1 if P-value > 0.05 and accept H_0 which means there is no relationship between gender and HIV Status of the event

4. Findings

Table 2: HIV Status Statistics.

HIV Status	Frequency	Percentage
Negative	30	30.6
Positive	65	66.3
Unknown	3	3.1
Total	98	100

Table 3: Weights of the Patients

Weight range (Kg)	Frequency	Percentage
1 - 20	3	3.0%
21 - 40	12	12.1%
41 - 60	52	52.9%
61 - 80	23	23.4%
81 - 100	6	6.1%
Above 100	2	2.0%

Table 4: Ages of the Patients.

Age group (Years)	Frequency	Percentage
0 – 16.7	12	12%
16.8 – 33.3	24	24.2%
33.4 - 50	44	44.3%
50.1 – 66.7	15	15.1%
66.8 – 83.3	1	1.0%
Above 83.3	2	2.0%

Table 5: 98 Tuberculosis Cases' Statistics.

The sex of patients	Total Number	No. of Events	No. of Censored	Censored Percentage
Male	50	45	5	10%
Female	48	42	6	12.5%
Overall	98	84	11	11.2%

Table 6: Means and Medians Survival Time

Sex of patients	Estimate	SE	mean		Estimate	SE	Median	
			95% C. I				95% C. I	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Male	28.685	2.231	24.311	33.059	26.600	1.970	22.739	30.461
Female	28.180	1.954	24.350	32.010	26.600	5.138	16.529	36.671
Overall	28.651	1.526	25.660	31.643	26.600	1.691	23.285	29.915

Table 7: Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	0.084	1	0.772

Table 8: Event Gender and HIV 2x2 Contingency Table

		HIV Status		TOTAL
		Negative	Positive	
GENDER	Male	11	34	45
	Female	13	26	39
TOTAL		24	60	84

Table 9: Event Chi-Square Table for Gender and HIV Status.

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	0.809	1	0.368		
Continuity Correction	0.432	1	0.511		
Likelihood Ratio	0.808	1	0.369		
Fisher's Exact Test				0.469	0.255
Linear-by-Linear Association	0.799	1	0.371		
No. of Valid Cases	84				

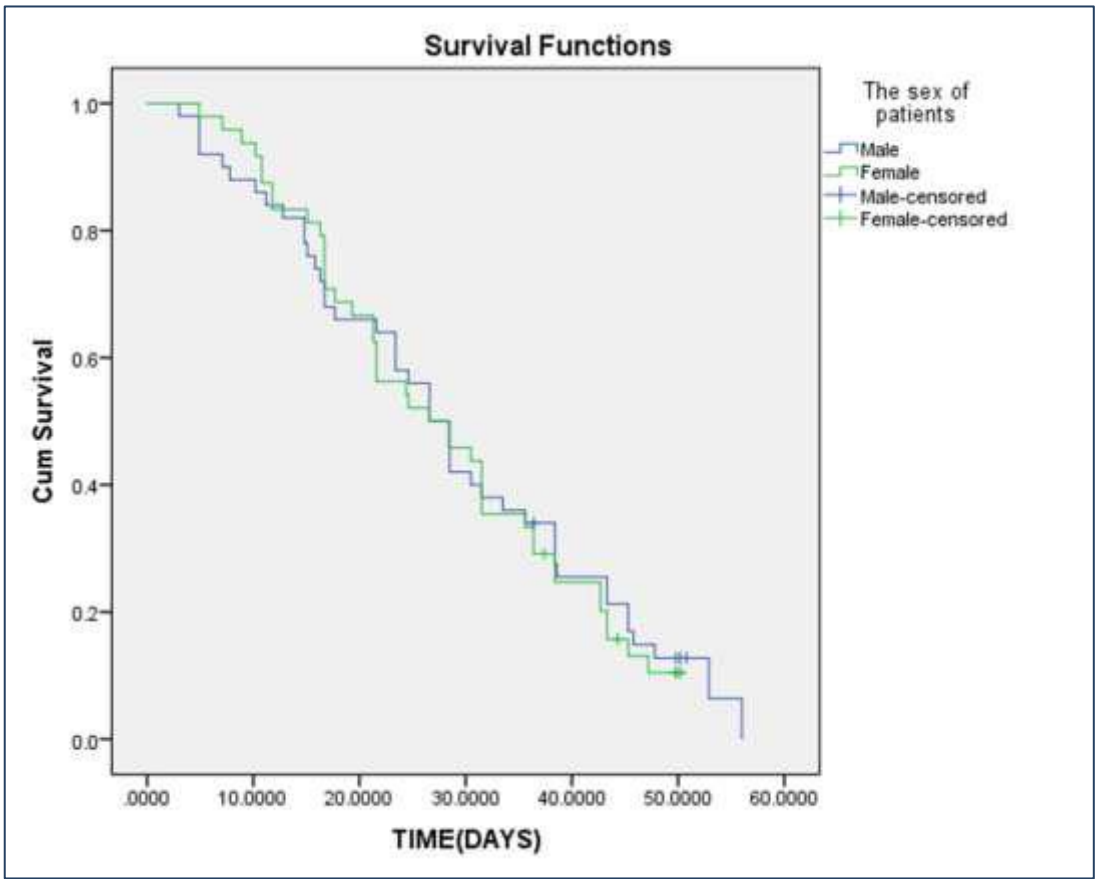

Figure 1: A graph of Cumulative Survival Function against Time in days.

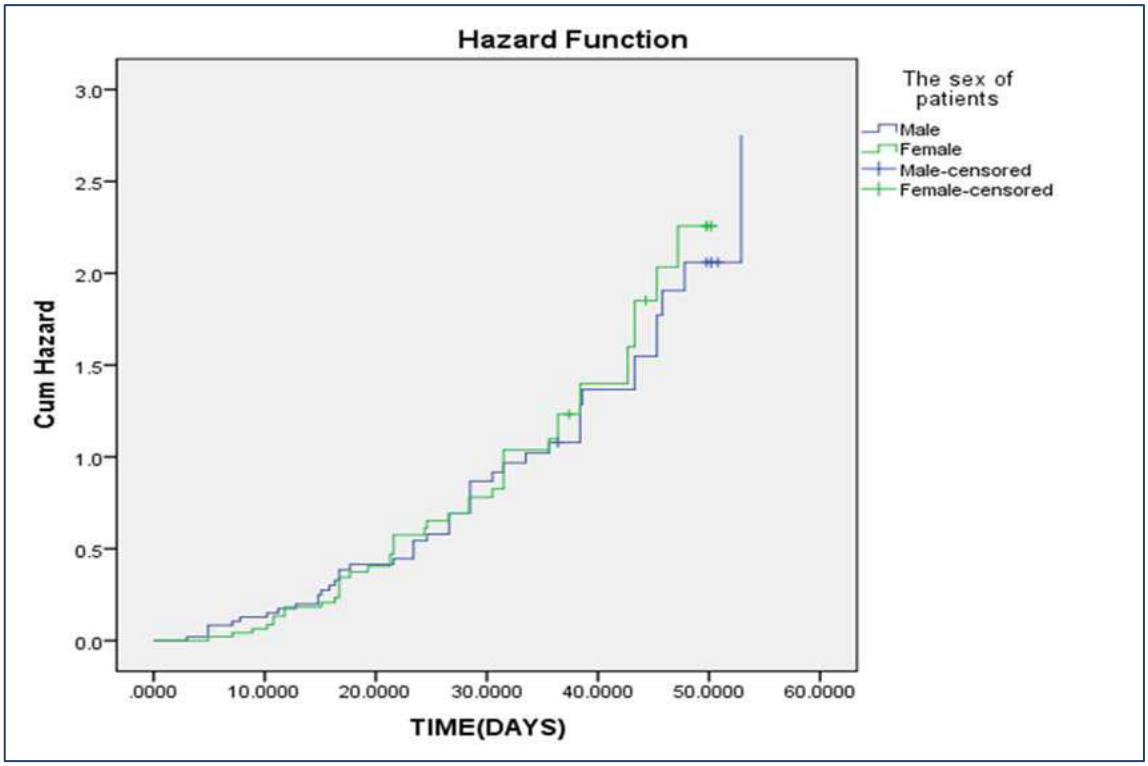

Figure 2: A graph of Cumulative Hazard Function against Time in days.

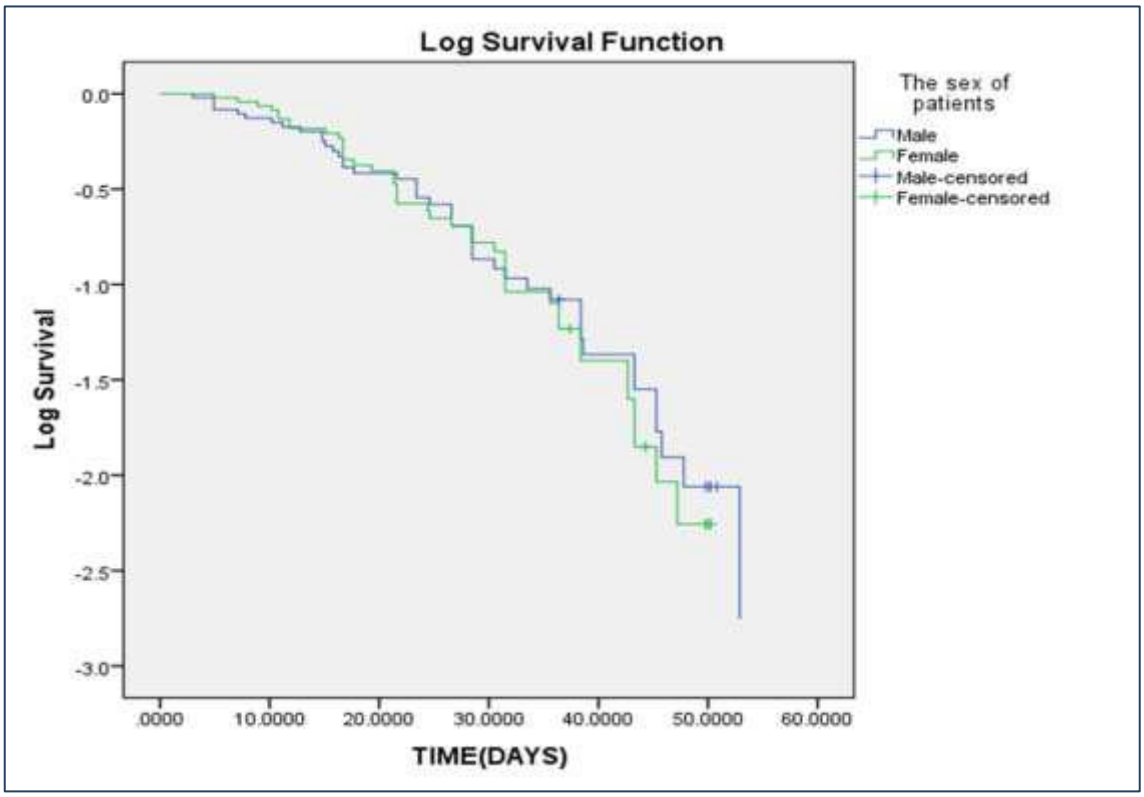

Figure 3: A graph of Log Survival Function against Time in days.

5. Conclusion and Recommendations

5.1 Conclusion

A total of 98 Tuberculosis patients were investigated in this study, with male patients forming 51.02% and female patients accounting for 48.98%; the total number of events that occurred were 87 (88.8%) and there were 11 (11.2%) censored cases. The mean survival time for the study was 28.651 days and the median survival time was 26.600 days, with mean survival time for male and female being 28.685 days and 28.180 days respectively, while the median survival time for both male and female was 26.600 days. At time 0, the survival probability was 1, and the probability of survival of about 31 days was 0.4, while the

probability of survival of more than 25 days for the entire cohorts was 50%. A log rank (Mantel-Cox) test showed $\chi^2 = 0.084$ and P-value = 0.772 > 0.05, indicating no sufficient evidence that there was difference between the male and female survival time. Out of the 98 cohorts, 65 (66.3%) were HIV positive, 30 (30.6%) were HIV negative and 3 (3.1%) were unknown, while 52.9% weighed between 41kg to 60kg and 44.3% were between 33 years and 50 years of age. Out of the cases that occurred, 53.57% were male and 71.43% were HIV positive. A Pearson Chi-Square test showed $\chi^2 = 0.809$ and P-value = 0.368 > 0.05, indicating no sufficient evidence that there was difference between gender and HIV status. In conclusion, the Kaplan Meier procedure was used to estimate the survival curves of Tuberculosis patients. The log-rank statistic was also used to test whether there is a significant difference in the survival experience of the Tuberculosis patients with respect to the different variables being considered. P < 0.05 was considered as the significance difference. The overall median survival time for Tuberculosis patients was 27 days. This indicated that 50% of the Tuberculosis patients survived longer than 27 days after being diagnosed with the dis ease. The log-rank test showed that the survival time between males and females was not significantly different at $\propto = 0.05$, and it was concluded that since most of the population survive pass 27 days, the treatment regimen is a good one.

5.2 Recommendations

Based on the results from this study, the following recommendations were made:

i. Further research could be carried out using Cox Proportional Hazard regression to incorporate more risk factors.
ii. A clinical study is recommended for research of this magnitude to capture more components and risk factors

References

[1] Abbara A, Kon O. M, Hansel. T, John. L, Davidson R. N, Sullivan. A, Barrett. J, Collin S. M, McGregor A. (2018), Time to diagnosis of tuberculosis is greater in older patients. *ERJ Open Resource journal* **5** (4): 00228.

[2] Ajagbe O. B, Kabair Z, O'Connor T. (2014), Survival analysis of adult tuberculosis disease. *PLoS ONE* **9** *(*11): 112838.

[3] Akessa, Geremew Muleta, Tadesse, Mulualem, Abebe, Gemeda (2015) Survival Analysis of Loss to Follow-Up Treatment among Tuberculosis Patients at Jimma University Specialized Hospital, Jimma, Southwest Ethiopia, *International Journal of Statistical Mechanics*, 2015, 923025, 7 pages

[4] Balaky S. T. J, Mawlood A. H, Shabila N. P. (2019), Survival analysis of patients with tuberculosis in Erbil, Iraqi Kurdistan region. *BMC Infect Dis*. **19** (1):865.

[5] Vincent Abiodun Michael & Ismaila Adewale Bolarinwa, 2020. "Parametric Survival Modeling of Tuberculosis Data- A Case Study of Federal Medical Centre, Bida, Nigeria," Modern Applied Science, Canadian Center of Science and Education, vol. 14(7), pages 1-37

[6] Carter B. B, Zhang. Y, Zou H. (2021), Survival analysis of patients with tuberculosis and risk factors for multidrug-resistant tuberculosis in Monrovia, Liberia. *PLoS One* **16** (4): 0249474.

[7] Kaplan, E. L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*. **53** (282): 457–481.

[8] Kleinbaum G. David and Mitchel Klein (2007). Statistics for Biology and Health Limenih and Workie. *BMC Public Health* 19:165

[9] Liao S., Cai C., F. M. Huo (2017), Trends in drug-resistant tuberculosis in China: data from a clinical tuberculosis Centre, *International Journal of Tuberculosis and Lung* **21** (9):990-995

[10] Liu K, Ai L, Pan J, Fei F, Chen S, Zhang Y, Wang W, Wu Q, Chen B, Pan J, Zhong J (2022), Survival Analysis and Associated Factors for Pulmonary Tuberculosis Death. Risk *Management Health policy* 15:1167-1178.

[11] Mulugeta G. (2018), Modeling the Survival of Tuberculosis Patients in Eastern region of Tigray Regional State, *Ethiopia;* 13: 473-481.

[12] Schober, P., & Vetter, T. R. (2018). Survival Analysis and Interpretation of Time-to-Event Data: The Tortoise and the Hare. *Anesthesia and analgesia*, *127*(3), 792–798

[13] Tekie D. K, Masango T. E, Nkosi Z. Z. (2018) Performance of the national tuberculosis control program in the post conflict Liberia. *PLoS ONE*. **13** (6): 0199474

[14] World Health Organization, Global Tuberculosis Report, WHO, Geneva, Switzerland, 2019.

[15] Yigzaw A. L. and Demeke L. W. (2018), Survival analysis of time to cure on multidrug resistance tuberculosis patients in Amhara region, Ethiopia. *BMC Public Health* 19: 165

[16] Zenebe T., Tefera E. (2020). Tuberculosis treatment outcome and associated factors among smear positive pulmonary tuberculosis patients in Afar, Eastern Ethiopia: a retrospective study. *Brazil Journal Infected Disease*. **20** (6):635–636.