Journal of Science Research and Reviews
PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2025.v2i1.27
Homepage: https://josrar.esrgngr.org
Original Research Article

Exploring Accelerated Failure Time Models for Tuberculosis Survival: Log-Logistic and Weibull Survival Regression Model

†[1]Usman, Abubakar, ‡[1]Doguwa, Sani I., §[1]Sadiq, Ibrahim A. and *[2]Akor, Augustina

[1]Department of Statistics, Faculty of Physical Sciences, Ahmadu Bello University, Zaria
[2]Doctrine Curriculum and Development Department, Armed Forces Command and Staff College Jaji, Kaduna State Nigeria.
*Corresponding Author's email: augustina.akor@gmail.com Phone: +2347067016940
ORDID iD: *https://orcid.org/0009-0004-2231-0832 †https://orcid.org/0000-0002-9104-1937
‡https://orcid.org/0000-0002-5779-2358 §https://orcid.org/0000-0002-2122-9344

KEYWORDS
Tuberculosis,
Survival Models,
AFT model,
Weibull Distribution,
Log-logistic Distribution,
Cox Proportional Hazard Models.

ABSTRACT
Tuberculosis (TB) remains a significant global health burden, necessitating robust statistical models to understand survival dynamics and inform interventions. Most survival analyses rely on the Cox Proportional Hazards (Cox PH) model, which may not adequately capture the survival time distribution. This study focuses on data from the National Tuberculosis and Leprosy Center (NTLC), Zaria, Kaduna State, Nigeria, to identify factors influencing TB survival and assess alternative parametric survival models. The study aims to identify factors associated with TB mortality, assess their impact on survival outcomes, and compare the performance of Weibull, and Log-Logistic Accelerated Failure Time (AFT) models to determine the most suitable model for TB survival data from NTLC Zaria. This study compares the performance of two Accelerated Failure Time (AFT) models, the Weibull, and the Log-Logistic in analyzing TB survival data. The analysis evaluates model fit using p-values, log-likelihood, and Akaike Information Criterion (AIC). Results indicate that the Weibull AFT model outperforms the Log-logistic, with the highest log-likelihood (-228.6) and the lowest AIC (485.11), and the Log-Logistic AFT model (AIC: 492.02, log-likelihood: -232.0). The p-values for both models demonstrate statistical significance, highlighting their effectiveness in modelling TB survival data. However, the Weibull model's higher performance suggests it better captures survival time variability in TB patients. These findings emphasize the importance of selecting appropriate survival models for TB data analysis and support the application of the Weibull AFT model for future studies. Further research should explore integrating advanced statistical techniques and machine learning approaches to enhance predictive accuracy and improve TB management strategies. This study contributes to this growing field by applying parametric survival models, to analyse TB survival data from the National Tuberculosis and Leprosy Centre (NTLC) in Zaria, Nigeria.

CITATION
Usman, A., Doguwa, S. I., Sadiq, I. A. &. Akor, A. (2025). Exploring Accelerated Failure Time Models for Tuberculosis Survival: Loglogistic and Weibull Survival Regression Model. *Journal of Science Research and Reviews*, *2*(1), 27-36. https://doi.org/10.70882/josrar.2025.v2i1.27

INTRODUCTION

According to Jurado *et al*. (2018), Tuberculosis (TB), caused primarily by *Mycobacterium tuberculosis*, is a highly infectious disease that mainly affects the lungs but can also impact other body organs. It is transmitted through airborne particles when an infected individual coughs, sneezes or speaks. TB is a leading global cause of mortality (WHO 2020), surpassing HIV in deaths, and often acts as an opportunistic disease in conjunction with immune-suppressive conditions like HIV/AIDS. TB is classified into Pulmonary TB (PTB), where the lungs are affected, and Extra-Pulmonary TB (EPTB), which involves other organs like the pleura, lymph nodes, and bones. PTB symptoms include persistent cough, chest pain, and fever, while EPTB symptoms vary based on the affected organ Xie *et al*. (2020).

Tuberculosis (TB) is a highly infectious airborne disease caused by various strains of mycobacteria. Despite being curable, TB remains one of the world's deadliest diseases and a leading cause of global morbidity and mortality, particularly in Africa and Asia, where poverty and inadequate healthcare systems exacerbate its impact. (WHO), 2009). In Nigeria, TB poses a significant public health challenge, with the country ranked fourth among the six nations accounting for 60% of the global TB burden. Many cases go unreported or undiagnosed, facilitating further disease spread.

The WHO's ten-year strategy to combat TB successfully reduced disease prevalence globally, saving millions of lives. However, progress in Nigeria has been slower, with over half a million new cases reported in 2014 alone, according to the National TB and Leprosy Control Programme (NTBLCP). Predictors of mortality among TB patients vary across regions, as do survival times and mortality rates. Africa, accounting for 12% of the global population, disproportionately bears 31% of the global TB burden. In 2015, the region's TB death rate stood at 46 per 100,000 people, the highest globally.

Treatment involves a six-month regimen of antibiotics, but drug-resistant strains such as Multidrug-Resistant TB (MDR-TB) and Extensively Drug-Resistant TB (XDR-TB) complicate management, requiring longer and more intensive therapies Okuonghae *et al*. (2007).

A retrospective cohort study by Mustapha *et al*. (2016) analyzed the survival times of tuberculosis (TB) patients using the Cox proportional hazards model. Data were obtained from TB/HIV cases registered between 2011 and 2013 at the University of Maiduguri Teaching Hospital (UMTH), Nigeria. The study found that HIV status significantly influenced the survival time of TB patients, while variables such as sex, age, residence, and TB type were not significant. However, the study did not explore other parametric PH models or Accelerated Failure Time (AFT) models, a gap this research aims to address.

Adamu *et al.,* (2017) analyzed a retrospective cohort of TB patients that commenced treatment between January 2010 and December 2014 in Aminu Kano Teaching Hospital. The Cox proportional hazards model was employed to determine risk factors for mortality. Results from their findings revealed that the risk factors for death were HIV-positive but not on anti-retroviral treatment, residence outside the city, previous TB treatment, no microbiological confirmation, having both pulmonary and extra-pulmonary TB and referral from a non-programme linked clinic/centre.

Xie *et al*. (2020) conducted a retrospective cohort study on 7,032 tuberculosis (TB) patients in the DOTS program in Tianjin, China, from 2014 to 2017, to identify risk factors for mortality during anti-TB treatment. They used the Kaplan–Meier method and Cox proportional hazards regression model for analysis. The study found that factors such as older age, male sex, HIV positivity, first sputum positivity, retreated TB, and patient delay (≥14 days) were associated with an increased risk of mortality during anti-TB treatment.

Okoro *et al*. (2022) studied the factors associated with treatment default among tuberculosis (TB) patients in Adamawa State, Nigeria, using the Cox proportional hazards regression and Kaplan-Meier method. The study assessed TB patient survival and treatment outcomes over six months (January to June 2019). The findings revealed that HIV-positive status, older age, primary education, and poor quality of life were significant risk factors for treatment default among TB patients in the region. Treatment default is identified as a key factor contributing to drug-resistant tuberculosis.

Many studies on tuberculosis (TB) data use Cox Proportional Hazards (PH) models without considering the assumptions that govern their use, which is often overlooked. This lack of focus on assumptions highlights the need for a more detailed analysis of the assumptions underlying these models. In some real data applications, Cox PH models may not provide a good fit, suggesting the need for alternative survival models. Accelerated failure-time models, which assume a parametric distribution for survival times, are especially useful in these cases, as they offer more accurate statistical inference and better model fitting (Swindell, 2009).

Survival analysis, a statistical methodology used to analyze time-to-event outcomes, provides a framework for studying the survival time and risk factors associated with patient outcomes Lee *et al*. (2003). This method contrasts with classical regression models like logistic and linear regressions, which cannot fully utilize incomplete data. Survival analysis focuses on the survival function, representing the likelihood of survival beyond a specific time point.

Other notable contributions to the development of parametric survival distributions include the introduction of the NOF-G family of distributions (Sadiq *et al.*, 2022), the NGOF-G family of distributions (Sadiq *et al.*, 2023a), and further extensions such as the NGOF-Et-G family (Sadiq *et al.*, 2023b) and the NGOF-OE-G family (Sadiq *et al.*, 2023c). Additionally, Obafemi *et al*. (2024) proposed the NETD distribution using a generalized logarithmic function, while Habu *et al*. (2024) advanced the field with an extension of the Topp-Leone (T-L) distribution, the regression model for diabetes risk factors (Sadiq *&* Komali, 2020), the general linear model for epilepsy (Sadiq *et al*., 2020). These advancements have significantly enriched the field of survival analysis by providing more flexible and robust models for capturing diverse data patterns. This study aims to identify factors associated with TB mortality, assess their impact on survival outcomes, and compare the performance of Weibull, and Log-Logistic Accelerated Failure Time (AFT) models to determine the most suitable model for TB survival data from NTLC Zaria.

MATERIALS AND METHODS

Method of Data Collection

This study is a retrospective cohort analysis of HIV/TB patients, using medical records from the National TB and Leprosy Center Hospital in Zaria, Kaduna State, from 2020 to 2023. The study aims to investigate survival time (duration until death) among these patients, with the dependent variable being survival time in months. The study focuses on identifying risk factors influencing survival, including gender, marital status, age, comorbidities, type of TB, smoking history, and alcohol use. Data collected is secondary, based on patient records, and includes censored observations for patients who did not die during the study period. A key advantage of this design is that it minimizes data collector bias by using historical data.

Data Analysis Framework

This study uses non-parametric methods, such as the Kaplan-Meier plot, to compare survival functions across different groups (Kaplan and Meier, 1958). A frequency distribution table summarizes the data collected from the National TB and Leprosy Center Hospital in Zaria, Kaduna State. Survival models, including parametric Proportional Hazards (PH), Accelerated Failure-Time (AFT), and semi-parametric models, will be used to identify risk factors influencing survival times in tuberculosis patients. The optimal model will be selected based on the lowest information criteria to ensure the best fit.

Kaplan-Meier (K-M) Estimator (1958) of the Survival Function

The Kaplan-Meier (KM) estimator is a non-parametric method used in survival analysis to estimate the survival function from time-to-event data (Kaplan and Meier, 1958). It is particularly valuable when dealing with censored data, where the event of interest has not occurred for all subjects within the study period. The Kaplan-Meier estimator also known as the product-limit estimator is the most widely used non-parametric method for estimating the survival function. The Kaplan-Meier estimator provides an estimate of the survival function $S(t)$, which represents the probability that an individual survives beyond time t:

$$S(t) = P(T > t) \quad (1)$$

where T is the random variable representing the time-to-event.

Then the K-M estimator of S(t) is defined:

The KM estimator is defined as:

$$\hat{S}(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right) \quad (2)$$

wheret_i: Time of the i^{th}event; d_i: Number of events (e.g., deaths, treatment completion) at t_i; n_i: Number of individuals at risk just before t_i and $\prod$: Product overall event times up to t.

The estimator adjusts for censored data by ensuring that only those individuals still at risk at time t are included in the calculation of the survival probability. It produces a step-function survival curve, which is widely used in practice.

Accelerated Failure Time Model (AFT)

The Accelerated Failure Time (AFT) model is a parametric survival analysis approach that directly models the time-to-event (e.g., failure or death). Unlike the proportional hazards model, which focuses on the hazard function, the AFT model describes how covariates accelerate or decelerate the time to an event.

The AFT model assumes that the survival time T is multiplied by a scaling factor due to the effect of covariates. This scaling factor accelerates or decelerates the event's occurrence, depending on the covariate values.

Model Formulation Accelerated Failure Time Model (AFT)

$$ln(T_i) = \boldsymbol{\beta}^T \boldsymbol{x}_i + \sigma\epsilon_i \quad (3)$$

where T_i: Survival time for individual i; $\boldsymbol{x}_i$: Vector of covariates for individual i; $\boldsymbol{\beta}$: Coefficient vector quantifying the effect of covariates; σ: Scale parameter and ϵ_i: Random error term with a specified distribution (e.g., exponential, Weibull, Log-logistic etc.).

Alternatively, the model in equation (3) can be written as:

$$T_i = T_0\, exp\{\boldsymbol{\beta}^T \boldsymbol{x}_i\} \quad (4)$$

where $T_0 = exp\{\sigma\epsilon_i\}$ represents the baseline survival time. Therefore, The survival functional form of the AFT model is:

$$S(t|\boldsymbol{x}_i) = S_0(t\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\}) \quad (5)$$

Log-Logistic Accelerated Failure Time (AFT) Model

The Log-Logistic Accelerated Failure Time (AFT) model is another commonly used parametric survival model. It is based on the log-logistic distribution, which is particularly useful when the hazard function exhibits non-monotonic behaviour (i.e., it increases initially and then decreases over time). The log-logistic distribution is characterized by its probability density function (PDF) and cumulative distribution function (CDF):

$$f(t) = \frac{\alpha\lambda^\alpha t^{\alpha-1}}{(1+(\lambda t)^\alpha)^2}, \quad t > 0 \quad (6)$$

where $\alpha > 0$: Shape parameter and $\lambda > 0$: Scale parameter (related to median survival time).

$$F(t) = \frac{(\lambda t)^\alpha}{1+(\lambda t)^\alpha} \quad (7)$$

The baseline survival function for the log-logistic distribution is given as:

$$S_0(t) = 1 - F(t) = \frac{1}{1+(\lambda t)^\alpha} \quad (8)$$

With covariates, however, by substituting equation (5) into equation (8), the log-logistic AFT model is as:

$$S(t|\boldsymbol{x}_i) = \frac{1}{1+(\lambda t\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\})^\alpha} \quad (9)$$

Likelihood Function of the Log-logistic AFT model

The likelihood function for the log-logistic AFT model is derived from the PDF and survival function. For n observations, the likelihood is:

$$L(\boldsymbol{\beta}, \sigma) = \prod_{i=1}^n f\ (t_i|\boldsymbol{x}_i)^{\delta_i} S(t_i|\boldsymbol{x}_i)^{1-\delta_i} \quad (10)$$

where $\delta_i = 1$ for uncensored observations (event occurred), $\delta_i = 0$ for censored observations (event not observed).

The log-likelihood function is:

$$ln\ L\ (\boldsymbol{\beta}, \sigma) = \sum_{i=1}^n [\delta_i\ ln\ f\ (t_i|\boldsymbol{x}_i) + (1-\delta_i)\ ln\ S\ (t_i|\boldsymbol{x}_i)] \quad (11)$$

Weibull Accelerated Failure Time Model (AFT)

The Weibull Accelerated Failure Time (AFT) model is an extension of the exponential AFT model, where survival times follow a Weibull distribution. Unlike the exponential distribution, the Weibull distribution allows for non-constant hazard rates, making it more flexible for modelling survival data. The Weibull AFT model assumes that the survival time T_0 in equation (4) follows a Weibull distribution with two parameters:

$$f(t) = \alpha\lambda t^{\alpha-1}\ exp\{-(\lambda t)^\alpha\}, \quad t > 0 \quad (12)$$

where $\alpha > 0$: Shape parameter, controlling the hazard rate and$\lambda > 0$: Scale parameter (related to the hazard rate).

Therefore, the baseline survival function for the Weibull distribution is given as:

$$S_0(t) = exp\{-(\lambda t)^\alpha\} \quad (13)$$

With covariates, however, by substituting equation (5) into equation (12), the Weibull AFT model is as:

$$S(t|\boldsymbol{x}_i) = S_0(t\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\}) = exp\{-(\lambda t\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\})^\alpha\} \quad (14)$$

Likelihood Function of the Weibull AFT Model

The likelihood function for the Weibull AFT model is derived from the Weibull distribution for n observations:

$$L(\alpha, \lambda, \boldsymbol{\beta}) = \prod_{i=1}^n \alpha\ \lambda^\alpha t_i^{\alpha-1}\ exp\{-(\lambda t_i\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\})^\alpha\} \quad (15)$$

The log-likelihood function is:

$$ln\ L\ (\alpha, \lambda, \boldsymbol{\beta}) = \sum_{i=1}^n [ln\ \alpha + \alpha\ ln\ \lambda + (\alpha-1)\ ln\ t_i - (\lambda t_i\ exp\{-\boldsymbol{\beta}^T\boldsymbol{x}_i\})^\alpha] \quad (16)$$

Parameters α, λ, and $\boldsymbol{\beta}$ in equation (15) can be estimated using maximum likelihood estimation (MLE).

RESULTS AND DISCUSSION

Description and Categorization of the covariates

The study incorporates several covariates, each categorized to facilitate analysis. Age groups of patients are classified as <35 years (0), 35–55 years (1), and >55 years (2). Body Mass Index (BMI) is categorized into Normal Weight (0), Underweight (1), and Overweight/Obesity (2). Gender is coded as Female (0) and Male (1), while marital status is classified as Single (0) and Married (1). The site of tuberculosis is noted as Yes (1) and No (0). Family history indicates the presence (1) or absence (0) of TB cases in the patient's family. Alcohol and smoking history are both recorded as Yes (1) and No (0). Comorbidities are categorized as No Comorbidity (0), HIV-AIDS (1), and Hepatitis (2). Initial tuberculosis treatment history is documented as No (0) and Yes (1). Lastly, the life status of patients is classified as Alive (Censored, 0) or Dead (1). Additional breakdown is given as follows:

Age: The age group of patients is categorized as follows:
0: <35 years
1: 35–55 years
2: >55 years

BMI (Kg/m^2): The Body Mass Index of patients is categorized into:
0: Normal Weight
1: Underweight
2: Overweight and Obesity

Gender: The sex of patients is coded as:
0: Female
1: Male

Marital Status: The marital status of patients is classified as:
0: Single
1: Married

Site of Tuberculosis: The type of tuberculosis (TB) associated with patients is recorded as:
1: Yes
0: No

Family History: Indicates whether there is a history of TB cases in the patient's family:
1: Yes
0: No
Alcohol History: Reflects the patient's history of alcohol consumption:
1: Yes
0: No
Smoking History: Indicates the patient's history of smoking:
1: Yes
0: No
Comorbidity: Represents the type of comorbidities experienced by the patient:
0: No Comorbidity
1: HIV-AIDS
2: Hepatitis
Initial TB Treatment History: Whether the patient had received initial tuberculosis treatment:
0: No
1: Yes
Life Status: Reflects the survival status of the patient:
0: Alive (Censored)
1: Dead

Distribution of Covariates in the Dataset

This provides an analysis of the demographic, clinical, and behavioural characteristics of a group of tuberculosis (TB) patients, along with their treatment history and outcomes. The demographic characteristics of tuberculosis (TB) patients reveal a male-dominated population, with 64% being male and 36% female. The majority of the patients fall within the middle age group of 35-55 years (59%), while younger patients under 35 years comprise 28%, and the elderly population over 55 years accounts for only 13%. This distribution suggests that TB disproportionately affects individuals in their prime working years, emphasizing the disease's socio-economic impact.

Clinically, the majority of TB cases are pulmonary (84%), with only 16% categorized as extra-pulmonary. Notably, 26% of patients reported a family history of TB and 74% did not, highlighting a significant hereditary or environmental exposure factor. Alcohol consumption is prevalent among these patients, with 76% admitting to a history of alcohol use and 24% not, smoking history is also significant, with 47% having a history of smoking and 53% not. These behavioural factors are likely contributing to disease susceptibility and progression.

Regarding comorbidities, the population is equally affected by hepatitis (36%) and HIV/AIDS (36%), while 28% reported no additional health conditions. This high prevalence of comorbidities emphasizes the vulnerability of TB patients to other infectious and chronic diseases, which can complicate treatment and outcomes. Moreover, a significant proportion of the patients were underweight (41%), with 24% overweight or obese, and 35% having a normal BMI. Malnutrition, as indicated by the underweight category, could be a contributing factor to weakened immunity and poorer prognosis.

In terms of treatment history and outcomes, 63% of the patients had received prior TB treatment, while 37% were new to treatment. The response variable represents the survival time to death (in months) with the censoring indicator showing that 80% of them were censored (either successfully treated or still undergoing treatment), whereas 20% resulted in death. This mortality rate emphasizes the seriousness of TB, particularly in the context of associated risk factors such as comorbidities, behavioural influences, and undernutrition. These findings highlight the need for integrated care approaches addressing both TB and its coexisting conditions to improve patient outcomes.

Kaplan-Meier Survival Analysis for the Tuberculosis Patients

For a more elaborate descriptive analysis, the study uses a non-parametric Kaplan-Meier approach to provide a summary of the distribution of the variables. The mean for the overall survival time of tuberculosis patients is 18.21 months. The maximum and minimum survival times for tuberculosis patients were 0 to 20 months. More than eighteen months is the probabilistic expected predetermined period, for which patients who exceed eighteen months after diagnosis are termed to have survived tuberculosis.

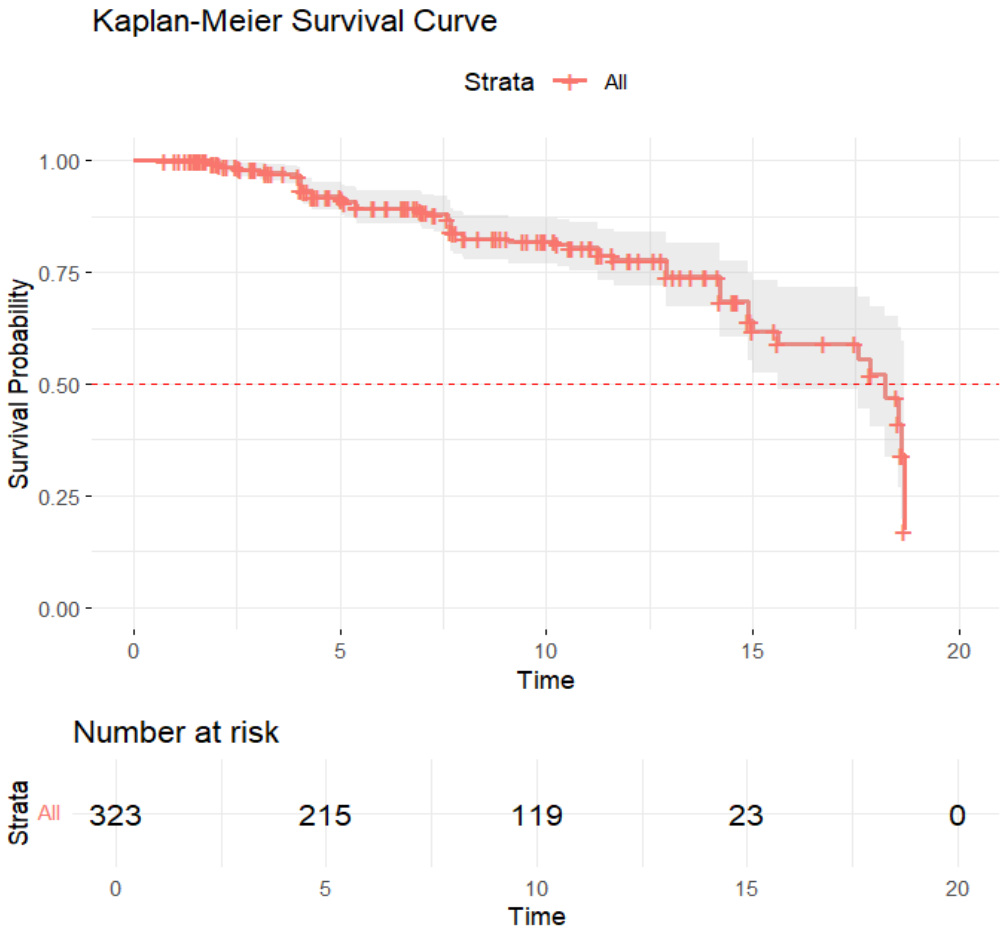

Figure 1: Kaplan-Meier probability of the survival time of Patients with Tuberculosis

Figure 1 depicts the Kaplan-Meier probability of the survival time of tuberculosis patients sampled with a 95 per cent confidence bound. The Kaplan-Meier uses the actual survival times and the probability of survival over time T.

Table 1: Kaplan-Meier Survival Probability of the TB Patients

Time	n_i	d_i	Survival Probability	Standard Error	95% Lower CI	95% Upper CI
1.840	302	2	0.993	0.005	0.984	1.000
2.100	296	1	0.990	0.006	0.979	1.000
2.140	293	1	0.987	0.007	0.974	1.000
2.530	286	1	0.983	0.007	0.969	0.998
2.600	284	1	0.980	0.008	0.964	0.996
3.060	276	1	0.976	0.009	0.959	0.994
3.220	274	1	0.973	0.010	0.954	0.992
3.680	267	1	0.969	0.010	0.949	0.989
3.980	266	1	0.965	0.011	0.944	0.987
4.040	262	8	0.936	0.015	0.908	0.965
4.110	238	1	0.932	0.015	0.903	0.962
4.340	235	3	0.920	0.016	0.888	0.953
5.030	214	1	0.916	0.017	0.883	0.949
5.060	212	1	0.911	0.017	0.878	0.946
5.130	209	1	0.907	0.018	0.873	0.943
5.360	207	1	0.903	0.018	0.868	0.939
5.390	206	1	0.898	0.019	0.862	0.936

5.420	203	1	0.894	0.019	0.857	0.932
6.940	185	1	0.889	0.020	0.851	0.928
7.000	182	1	0.884	0.020	0.846	0.924
7.270	178	1	0.879	0.021	0.840	0.920
7.590	174	2	0.869	0.022	0.828	0.912
7.690	167	5	0.843	0.024	0.798	0.891
7.760	149	1	0.837	0.024	0.791	0.886
7.960	145	1	0.832	0.025	0.784	0.882
7.990	144	1	0.826	0.025	0.778	0.877
9.070	131	1	0.820	0.026	0.770	0.872
10.260	117	1	0.813	0.027	0.762	0.866
10.550	115	1	0.805	0.027	0.754	0.861
11.240	91	2	0.788	0.029	0.732	0.848
11.640	72	1	0.777	0.031	0.718	0.840
12.920	64	3	0.740	0.036	0.673	0.814
14.200	53	4	0.685	0.043	0.606	0.774
14.890	32	2	0.642	0.050	0.552	0.747
14.990	28	1	0.619	0.053	0.523	0.732
15.620	22	1	0.591	0.057	0.488	0.715
17.560	17	1	0.556	0.064	0.444	0.696
17.850	16	1	0.521	0.069	0.403	0.675
18.210	10	1	0.469	0.079	0.337	0.653
18.540	8	1	0.410	0.088	0.269	0.626
18.610	6	1	0.342	0.097	0.197	0.595
18.670	2	1	0.171	0.130	0.039	0.760

Table 1 provides insights into the survival probabilities of tuberculosis (TB) patients over time using Kaplan-Meier estimates. At the earliest recorded time (1.840 months), the survival probability is high at 0.993 (99.3%), indicating that most patients are still alive or event-free at this point. Over time, the survival probability decreases steadily, reflecting the occurrence of deaths or events among the patients.

At this time (4.04 months), there is a noticeable decline in survival probability from 0.965 to 0.936, likely due to the occurrence of 8 events. This sharp drop indicates a critical point where many patients experienced the event of interest. Significant decreases in survival probabilities are also observed at 7.69 months (5 events) and 14.2 months (4 events), suggesting clusters of adverse outcomes.

By the 18.67-month mark, the survival probability drops to 0.17 (17.1%), indicating that only a small fraction of patients remain event-free by this time. The wide confidence intervals (95% CI of 0.039 to 0.760) at later times suggest uncertainty due to a smaller number of patients remaining under observation.

Early in the study, the 95% confidence intervals (e.g., 0.984 to 1.000 at 1.840 months) are narrow, reflecting high precision due to a larger sample size. As time progresses and fewer patients remain at risk, the confidence intervals widen (e.g., 0.039 to 0.760 at 18.670 months), indicating reduced precision in survival probability estimates.

The Kaplan-Meier survival curve shows a gradual decline in survival probability, highlighting the progressive nature of TB's impact on patient survival. Critical points in time (e.g., 4.04, 7.69, and 14.2 months) suggest periods where interventions or enhanced monitoring might be necessary to improve patient outcomes. The early and steep declines suggest the importance of timely diagnosis and treatment to improve survival outcomes. This analysis emphasizes the practical use of Kaplan-Meier estimates in understanding survival trends and identifying critical periods for intervention in TB management.

Table 2: Log-logistic AFT Model fitted to the Tuberculosis (TB) Dataset

Covariates	Coefficients	Standard Error	z	p
(Intercept)	2.3484	0.3986	5.89	3.80e-09
as.factor(Age)> 55 years	-0.7817	0.2505	-3.12	0.0018
as.factor(Age)35-55 years	-0.4628	0.2234	-2.07	0.0383
as.factor(Gender)Male	-0.2935	0.1818	-1.61	0.1063
as.factor(SiteTB)Pulmonary	0.7555	0.1581	4.78	1.80e-06
as.factor(FamilyHist)Yes	-0.3932	0.1465	-2.68	0.0073

as.factor(AlcoholHist)Yes	0.51800	0.1819	2.85	0.0044
as.factor(SmokingHist)Yes	-0.4650	0.1555	-2.99	0.0028
as.factor(Comorbodity)HIV-AIDS	0.2250	0.1634	1.38	0.1686
as.factor(Comorbodity)No comorbidity	0.4207	0.2289	1.84	0.0661
as.factor(TB treatment)Yes	0.2893	0.1510	1.92	0.0554
as.factor(BMI)Overweight and Obesity	0.2320	0.2071	1.12	0.2627
as.factor(BMI)Under Weight	0.4360	0.1691	2.58	0.0099
Log(scale)	-0.9616	0.0981	-9.8	< 2e-16
Models	χ^2	p-value	Loglik	AIC
Log-logistic AFT	112.84	1.6e-18	-232.0	492.0243

Table 2 presents the results of the Log-logistic AFT model fitted to the tuberculosis (TB) dataset. The model's log-likelihood is -232.0, significantly improved from the intercept-only model's -288.4, indicating that the inclusion of covariates improves the fit. Then $(\chi^2 = 112.84, df = 12, p = 1.6 \times 10^{-18})$, showing that the covariates collectively contribute significantly to the model. The Akaike Information Criterion (AIC) is 492.02, useful for comparing this model with others. The scale parameter (0.382) indicates the dispersion of survival times. A smaller scale suggests a narrower distribution of survival times, characteristic of the log-logistic model.

Patients aged >55 years $(z = -3.12, p = 0.0018)$ and 35–55 years $(z = -2.07, p = 0.0383)$ have significantly shorter survival times compared to those aged <35 years. Male patients $(p = 0.1063)$ show no significant difference in survival compared to females. Pulmonary TB $(z = 4.78, p = 1.8 \times 10^{-6})$ is associated with significantly longer survival times compared to extrapulmonary TB. Having a family history of TB $(z = -2.68, p = 0.0073)$ is associated with significantly shorter survival times. Patients with a history of alcohol consumption $(z = 2.85, p = 0.0044)$ have significantly longer survival times. A history of smoking $(z = -2.99, p = 0.0028)$ is associated with shorter survival times. HIV/AIDS $(p = 0.1686)$ and "No comorbidity" $(p = 0.0661)$ are not significant predictors of survival.

TB treatment $(z = 1.92, p = 0.0554)$ shows borderline significance, suggesting a potential association with longer survival. Underweight patients $(z = 2.58, p = 0.0099)$have significantly longer survival while being overweight or obese $(p = 0.2627)$does not significantly affect survival.

Age, TB site, family history, alcohol history, smoking history, and being underweight significantly influence survival time. Gender, comorbidities, and being overweight/obese do not significantly affect survival. TB treatment $(z = 1.92, p = 0.0554)$ approaches significance, warranting further investigation.

Focus on older patients, smokers, and those with a family history of TB for targeted interventions to improve survival. While TB treatment shows borderline significance, further investigation with larger sample sizes may clarify its role in improving survival.

Table 3: Weibull AFT Model fitted to the Tuberculosis (TB) Dataset

Covariates	Coefficients	Standard Error	z	p
(Intercept)	2.8114	0.3739	7.5200	5.50e-14
as.factor(Age)> 55 years	-0.7649	0.2294	-3.3300	0.00086
as.factor(Age)35-55 years	-0.5082	0.2029	-2.500	0.01228
as.factor(Gender)Male	-0.2853	0.1843	-1.5500	0.12166
as.factor(SiteTB)Pulmonary	0.5852	0.1315	4.4500	8.60e-06
as.factor(FamilyHist)Yes	-0.4787	0.1324	-3.6100	0.00030
as.factor(AlcoholHist)Yes	0.4207	0.1667	2.5200	0.01161
as.factor(SmokingHist)Yes	-0.4981	0.1402	-3.5500	0.00038
as.factor(Comorbodity)HIV-AIDS	0.0857	0.1475	0.5800	0.56131
as.factor(Comorbodity)No comorbidity	0.2547	0.2129	1.2000	0.23141
as.factor(TB treatment)Yes	0.3544	0.1314	2.7000	0.00700
as.factor(BMI)Overweight and Obesity	0.2272	0.1832	1.2400	0.21483
as.factor(BMI)Under Weight	0.5061	0.1446	3.5000	0.00046
Log(scale)	-0.8047	0.0981	-8.2100	2.30e-16
Models	χ^2	p-value	Loglik	AIC
Weibull AFT	117.46	2.0e-19	-228.6	485.1086

Table 3 presents the results of the Weibull AFT model fitted to the tuberculosis (TB) dataset. The model's log-likelihood is -228.6, significantly improved from the intercept-only model's -287.3, indicating that the inclusion of covariates improves the model's fit. Then ($\chi^2 = 117.46, df = 12, p = 2 \times 10^{-19}$), demonstrates that the covariates collectively contribute significantly to the model. The Akaike Information Criterion (AIC) is 485.11, useful for model comparison. A Weibull scale parameter less than 1 suggests decreasing hazard rates over time.

Patients aged >55 years ($z = -3.33, p = 0.00086$) and 35-55 years ($z = -2.50, p = 0.01228$) have significantly shorter survival times compared to those aged <35 years. Male patients ($p = 0.12166$) show no significant difference in survival compared to females. Pulmonary TB ($z = 4.45, p = 8.6 \times 10^{-6}$) is associated with longer survival compared to extrapulmonary TB. Having a family history of TB ($z = -3.61, p = 0.00030$) is associated with shorter survival. Patients with a history of alcohol consumption ($z = 2.52, p = 0.01161$) have significantly longer survival. A history of smoking ($z = -3.55, p = 0.00038$)is associated with shorter survival. Neither HIV/AIDS ($p = 0.56131$)nor "No comorbidity" ($p = 0.23141$) are significantly associated with survival. Patients receiving TB treatment ($z = 2.70, p = 0.00700$)have significantly longer survival. Being underweight ($z = 3.50, p = 0.00046$)is associated with longer survival, but being overweight or obese ($p = 0.21483$) shows no significant effect.

Age, TB site, family history, alcohol history, smoking history, TB treatment, and being underweight are significant predictors of survival time. Gender, comorbidities, and being overweight/obese do not significantly affect survival time. The decreasing hazard rate of the Weibull AFT model suggests that the risk of death declines as survival time increases, which aligns with the disease progression in treated TB patients. These findings emphasize the importance of targeted interventions for older patients, smokers, and those with a family history of TB to improve survival outcomes.

Table 4: AFT Models Performance Comparison fitted to the Tuberculosis (TB) Dataset

Models	χ^2	p-value	Loglikelihood	AIC
Log-logistic AFT	112.84	1.6e-18	-232.0	492.0243
Weibull AFT	117.46	2.0e-19	-228.6	485.1086

Table 4 presents the comparative performance of three parametric Accelerated Failure Time (AFT) models, Log-logistic and Weibull fitted to the tuberculosis (TB) dataset. The performance metrics include the chi-square statistic, p-value, log-likelihood, and Akaike Information Criterion (AIC). All models have statistically significant chi-square values ($p < 0.05$), indicating that the covariates contribute meaningfully to explaining the survival times. Among the models, the Weibull AFT model has the highest chi-square statistic (117.46), suggesting the strongest explanatory power for survival time variability. The Weibull AFT model has the highest log-likelihood (-228.6), indicating the best fit to the data among the two models. The AIC measures model quality, balancing fit and complexity. Lower AIC values indicate better models. The Weibull AFT model has the lowest AIC (485.1086), demonstrating the best trade-off between model fit and parsimony.

The Log-logistic AFT model is simplistic with constant hazard assumption and has the lowest chi-square value (99.45) and highest AIC (492.0243), indicating it performs poorly compared to the Weibull model. The Weibull AFT model is flexible with varying hazard rates. Best overall performance, as it has the highest chi-square value, highest log-likelihood, and lowest AIC. This model is most suitable for the TB dataset. The Weibull AFT model is the best-fitting model for the tuberculosis dataset, as it provides the best balance of explanatory power and model parsimony. It is recommended for use in further analysis and interpretation of TB survival data.

CONCLUSION

This study highlights the importance of parametric survival models in understanding the survival patterns and risk factors influencing tuberculosis (TB) outcomes. The Weibull AFT model emerged as the most suitable for analyzing TB survival data from the National Tuberculosis and Leprosy Center (NTLC) in Zaria, Nigeria, outperforming other models in capturing the time-to-event dynamics. Key risk factors, including age, TB site, smoking history, and body mass index, were identified as significant determinants of mortality. The findings underscore the need for tailored interventions targeting these risk factors to improve patient survival. Additionally, the study emphasizes the value of applying diverse survival models to ensure robust and accurate analyses, contributing to evidence-based strategies for TB management in Nigeria. The policy and intervention strategies, targeted interventions should focus on addressing significant risk factors such as smoking, underweight status, and advanced age to improve TB survival outcomes. The patient-centred care, efforts should be made to enhance nutritional support for TB patients, especially those underweight, and provide additional care for older patients and those with pulmonary TB. A broader model applications, researchers are encouraged to explore

diverse survival models beyond Cox PH to gain deeper insights into the survival distributions of TB patients in various contexts. Strengthen awareness campaigns focusing on the risks associated with smoking and poor nutritional status to mitigate their impact on TB progression and mortality.

REFERENCES

Adamu, A. L., Gadanya, M. A., Abubakar, I. S., Jibo, A. M., Bello, M. M., Gajida, A. U., ... & Abubakar, I. (2017). High Mortality Among Tuberculosis Patients on Treatment in Nigeria: A Retrospective Cohort Study. *BMC Infectious Diseases*, *17*, 1-11.

Cox, D. R. (2018). *Analysis of Survival data*. Chapman and Hall/CRC.

Habu, L., Usman, A., Sadiq, I. A., & Abdullahi, U. A. (2024). Estimation of Extension of Topp- Leone Distribution using Two Different Methods: Maximum Product Spacing and Maximum Likelihood Estimate. *UMYU Scientifica*, *3*(2), 133-138.

Jurado, L. F., & Palacios, D. M. (2018). Tuberculosis: A Risk Factor Approach. *Tuberculosis*, *16*, 37-52.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, *53*(282), 457-481.

Lee, E. T., & Wang, J. W. (2003). Some well-known Parametric Survival Distributions and their Applications. *Statistical Methods for Survival Data Analysis*, 134-161.

Mustapha, M., Bakari, H. R., & Lawal, A. (2016). *Survival Times of Tuberculosis Patients Using Cox Proportional Hazards*. *2*(8), 3–14.

Obafemi, A. A., Usman, A., Sadiq, I. A., & Okon, U. (2024). A New Extension of Topp-Leone Distribution (NETD) Using Generalized Logarithmic Function. *UMYU Scientifica*, *3*(4), 127-133.

Okoro, C., Bamigbala, O. A., Ojetunde, A. O., & Ibrahim, A. (2022). Risk Factors Associated with Treatment Default Among Tuberculosis Patients in Adamawa State, Nigeria. *Galician Medical Journal*, *29*(2), E202221-E202221.

Okuonghae, D., & Korobeinikov, A. (2007). Dynamics of Tuberculosis: The Effect of Direct Observation Therapy Strategy (DOTS) in Nigeria. *Mathematical Modelling of Natural Phenomena*, *2*(1), 113-128.

Sadiq, I. A., & Komali, K. (2020). On the Study of Clear Causal Risk Factors of Diabetes Mellitus Using Multiple Regression. *Asian Journal of Probability and Statistics*, *9*(4), 29-47.

Sadiq, I. A., Doguwa, S. I. S., Yahaya, A., & Garba, J. (2023c). New Generalized Odd Fréchet-Odd Exponential-G Family of Distribution with Statistical Properties and Applications. *FUDMA Journal of Sciences*, *7*(6), 41-51.

Sadiq, I. A., Doguwa, S. I. S., Yahaya, A., & Usman, A. (2023b). Development of New Generalized Odd Fréchet-Exponentiated-G Family of Distribution. *UMYU Scientifica*, *2*(4), 169-178.

Sadiq, I. A., Doguwa, S. I., Yahaya, A., & Garba, J. (2022). New Odd Frechet-G Family of Distribution with Statistical Properties and Applications. *AFIT Journal of Science and Engineering Research,* 2(2), 84-103.

Sadiq, I. A., Doguwa, S. I., Yahaya, A., & Garba, J. (2023a). New Generalized Odd Frechet-G (NGOF-G) Family of Distribution with Statistical Properties and Applications. *UMYU Scientifica*, *2*(3), 100-107.

Sadiq, I. A., Raghav, J. S., & Sharma, S. K. (2020). Conglomeration of General Linear Model for Epilepsy Clinical Neuroimaging. *Asian Journal of Probability and Statistics*, *10*(2), 1-12.

Swindell, W. R. (2009). Accelerated failure time models provide a useful statistical framework for aging research. *Experimental gerontology*, *44*(3), 190-200.

Xie, Y., Han, J., Yu, W., Wu, J., Li, X., & Chen, H. (2020). Survival Analysis of Risk Factors for Mortality in a Cohort of Patients with Tuberculosis. *Canadian Respiratory Journal*, *2020*(1), 1654653.