

Survival Analysis of Neonatal Jaundice Patients in Ahmadu Bello University Teaching Hospital, Shika Zaria

Enesi Latifat O.^{1,*}, Turkur D.², Dikko H. G.¹

¹Department of Statistics, Ahmadu Bello University, Zaria

²Department of Community Medicine, Ahmadu Bello University, Zaria

Abstract In this work, Survival analysis of neonatal jaundice patients using Accelerated Failure Time(AFT) Model was carried out. The parametric estimates of the model were derived using maximum likelihood method. Furthermore, we applied the parametric model to a secondary data of 150 neonatal jaundice patients collected from Ahmadu Bello University Teaching Hospital, Shika, Zaria. The Predictive models that could be used to predict the survival times of neonatal jaundice patients were formulated. The data was analysed using R statistical software to estimate the parameters. Kaplan Meier (KM) curves were plotted for the seven categorical variables with time recorded in days indicated that patients who use tradition medicine have a higher survival time than patients who employed orthodox. Patients who do not visit Antenatal care have shorter survival time than those who used the Antenatal care. The results for the GWL AFT model shows that the survival time of neonates increased statistically and significantly because they were influenced by Antenatal visit, weight, type of delivery and gender respectively. The parametric model of the Generalized Lindley-Weibull (GLW) model performed better when compared with the results of some common existing parametric survival models based on the performance measures.

Keywords Generalized Lindley-Weibull (GLW), Kaplan Meier, Neonatal jaundice

1. Introduction

One of the features of survival data structure is censoring. Survival data are generally collected for a time interval in which the occurrences of a specific event are obtained. Survival data modelling has plays a vital roles in the application of statistics in epidemiology [2], thus, survival analysis of neonatal jaundice patient was carried out to model time to surviving neonatal jaundice. In survival analysis, censoring is the loss of observation on the lifetime variable of interest in the middle of the study due to various reasons. For most censoring types, a section of survival time for censored observations are observable and can be utilized in calculating the risk of experiencing a particular event. Censoring is generally divided into several types, such as right, left and interval censoring. Right censoring is the type of censoring that occurred when an individual has entered a study but is lost to follow-up, the actual event time is placed somewhere to the right of the censored time along the time axis [13]. As right censoring occurs more frequent than other types and its information cannot be excluded in the estimation of a survival model. Left censoring can occur when an individual failed before data collection began.

Interval censoring occurs when an individual requires follow-ups or inspections. The interval is, begin at zero or end at infinity.

Cox Proportional Hazard model is also referred to as semi-parametric model. It makes assumption about parametric form for the effect of the predictors on the hazard but no assumptions about non-parametric model being make. The interpretation of Cox proportional or Cox regression model means that the effects of different variables on survival are constant over time. One of the basic assumptions in a Cox regression model is the proportional hazard assumption, since it does not assume any form of distribution for survival times [3].

Proportional hazard simply means that the ratio of the hazards for two subjects, say, O and P at time t is the same for all values of t. Also, if the assumption of proportional hazards is true for two groups of survival times, say Group O and Group P, then the true survivor functions for the two groups fail to cross each other then we say they are parallel which is necessary but not a sufficient condition for proportional hazards. In modeling with explanatory variables, assumption of proportional hazards means that the effects of the different variables on survival are constant over time [7].

Parametric model can be estimated either by the proportional hazard or the Accelerated Failure Time (AFT) regression model. In this study, we will adopt AFT Model.

* Corresponding author:

Beautylatienes@yahoo.com (Enesi Latifat O.)

Received: Dec. 30, 2021; Accepted: Feb. 16, 2022; Published: Apr. 15, 2022

Published online at <http://journal.sapub.org/statistics>

The AFT model describes a situation where the mechanical or biological life history of an event is accelerated (or decelerated). One advantage of using the specification of AFT regression model is that it covers a wide scope of the survival time distribution [13]. AFT Model is a general model for data consisting of survival times, in which explanatory variables measured on an individual are assumed to act multiplicatively on the time dependent, and so affect the rate at which an individual proceeds along the time axis. Consequently, the model can be interpreted in terms of the speed of progression of a disease, for example, those receiving treatment say, A and those receiving treatment say, B, This model assumes that the survival time of an individual on one treatment is a multiple of the survival time on the other treatment; as a result, the probability that an individual on treatment A survives beyond time t is the probability that an individual on treatment B survives beyond time [5].

An important feature of the new families of generalized distributions is based on their flexibility and compatibility to analysed real life data [2]. Over the period, Generalized Weibull-Lindley (GLW) distribution has been developed by many researcher to model survival data such as [6], [12], their study GLW appeared to be more flexible when compared with other distribution. Jaundice is the most common disease in infants within the first week of life and it causes an infant mortality if not diagnosed and treated on time. The preventable factors that cause the readmission of the infant after birth is very important. This inspires the study of survival of neonatal jaundice with possibly time dependent covariates. Therefore, we are looking for the best AFT model to predict the most common factors rising neonatal jaundice. And there is also need to develop the Generalized Lindley-weibull Survival Model in order to investigate some associated factors to the survival time of the neonatal jaundice patients.

2. Materials and Methods

The survival analysis of time to event data however covers a wide scope which involves several parametric distributions as could be possible, but the researcher has narrowed down the study to cover four distributions. For the purpose of this research, application of the AFT models (Exponential, Weibull, Log-logistics, and Generalized Lindely -Weibull) in analyzing neonatal jaundice data will be discussed.

2.1. Survival Function

The survival function is a probability of an individual surviving longer than a given period of time t [4]. Let T denote a continuous non-negative random variable representing time until some event of interest, with probability density function (pdf) $f(t)$ and cumulative distribution function (cdf) $F(t) = Pr\{T \leq t\}$, the probability of being alive at t . The survival function is given by $S(t)$ and

$$S(t) = P(\text{an individual survives longer than } t)$$

$$P(T > t)$$

$$S(t) = 1 - P(T \leq t) \quad (1)$$

using the definition of cumulative distribution function $F(t)$ of T , we can write;

$$S(t) = 1 - F(t) \quad (2)$$

Here $S(t)$ is a non-increasing function of time t with the following properties. (i) i.e. at the beginning of the study (ii) $S(t) = 1$ for $t = 0$ $S(t) = 0$ for $t \rightarrow \infty$, which means that everyone will eventually experience the event. However, we will also allow the possibility that $S(\infty) > 0$. This corresponds to a situation where there is a non-negative probability of surviving or not experiencing event. For example, if the event of interest is the time from response until the disease relapse and the disease has a cure for some proportion of individuals in the population, then we have $S(\infty) = 0$, $S(\infty)$ corresponds to the proportion of cured individuals.

2.2. Hazard Function

The hazard function $h(t)$ is also referred to as the conditional mortality rate or instantaneous failure rate or force of mortality. This represents the conditional mortality rate that is, event has not yet occurred [10].

$h(t)$ in terms of cumulative distribution function $F(t)$ and probability density function $f(t)$ is given by

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (3)$$

$$h(t) = \frac{f(t)}{S(t)} \quad (4)$$

The hazard function describes the concept of the risk of an outcome (e.g., failure, death) in an interval after time t , conditional on the subject having survived time t . The hazard functions seems to be more intuitive to use in survival analysis than the pdf because it attempts to quantify the instantaneous risk that an event will take place at time t given that the subject survived to time t . A number of models are available for the analysis of the relationship of a set of predictor variables with the survival time.

2.3. Non Parametric Test

2.3.1. Kaplan-Meier Estimate of the Survival Function

Suppose that k individuals have died in a group of individuals. Let $0 \leq t_{(1)} < t_{(2)} < \dots < t_{(k)} < \infty$ be the observed ordered death times. Let k_j be the size of the risk set at $t_{(j)}$, where risk set denotes the collection of individuals alive and uncensored just before $t_{(j)}$. Let $C_{(j)}$ be the number of observed deaths at $t_{(j)}$, $j = 1, 2, \dots, k$. Then the

Kaplan-Meier(K-M) estimator of $S(t)$ is given as

$$\hat{S}(t) = \prod_{j:t(j) \leq t} (1 - \frac{c_j}{k_j}) \quad (5)$$

This estimator is a step function that changes values only at the time of each death. In fact, K-M estimator will be shown next to maximize the likelihood in the discrete case.

Let $e_{i=j}$ if the i th individual is censored at k_j . Then, the log likelihood contribution to the total likelihood is

$$\sum_{k < e_i} \log(1 - p_k) \quad (6)$$

Then the total log likelihood is given by

$$I = \sum_{deat h_i} \log h_{bi} + \sum_{deat h_i} [\sum_{k < b_i} \log(1 - p_k)] + \sum_{censor i} \sum_{k < e_i} \log(1 - p_k) \quad (7)$$

$$= \sum_{deat h_i} b_j \log h_j + \sum_{deat h_i} [\sum_{k < j} d_j] \log(1 - p_k) + \sum_{k [\sum_{k \leq j} c_j] \log(1 - p_k) \\ = \sum_{k < e_i} [b_j \log p_j + (k_j - b_j) \log(1 - p_j)] \quad (8)$$

Where b_j the number of observed death at is k_j , c_j is the number censored at (k_j, k_{j+1}) and k_j is the number of living and uncensored at k_{j+1}

If p_j is the solution of

$$\frac{\delta y}{\delta p_j} = \frac{b_j}{p_j} - \frac{k_j - b_j}{1 - p_j} = 0, \quad (3.11)$$

then

$$\hat{p}_j = \frac{b_j}{k_j} \quad (*)$$

This maximizes the likelihood since the total log likelihood function is concave down. So that the K-M estimator of the survival function is

$$\hat{S}(t) = \prod_{j|k(j) \leq t} (1 - \hat{p}_j) \quad (9)$$

Substitute equation (*) into equation (9), we have

$$= \prod_{j|k(j) \leq t} (1 - \frac{b_j}{k_j}) \quad (10)$$

Therefore, the K-M estimator is the maximum likelihood estimator. The K-M estimator gives a discrete distribution, if the observations are modelled to come from unknown continuous distribution, the maximum likelihood estimator does not exist [8].

2.4. Cox Proportional Hazards

The Cox Proportional Hazards model is proposed by [3]

$$h(t|X) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_r x_r) \\ = h_0(t) \exp(\beta' x) \quad (11)$$

where $h_0(t)$ is called the baseline hazard function, which is the hazard function for an individual for whom all the variables included in the model are zero, $X = (x_1, x_2, \dots, x_r)'$ is the values of the vector of explanatory variables for a particular individual, and $\beta' = (\beta_1, \beta_2, \dots, \beta_r)$ is a vector of regression coefficients. The corresponding survival functions are related as follows:

$$S(t|X) = S_0(t) \exp \sum_{i=1}^r \beta_i x_i \quad (12)$$

This model is known as the Cox regression model, it makes no assumptions about the form of $h_0(t)$ (non-parametric part of model) but assumes parametric form for the effect of the predictors on the hazard (parametric part of model). The model is therefore referred to as semi-parametric model. The measure of effect is called hazard ratio. The hazard ratio of two individuals with different covariates X and X^* is

$$\widehat{HR} = \frac{h_0(t) \exp(\hat{\beta}' X)}{h_0(t) \exp(\hat{\beta}' X^*)} = \exp(\sum \hat{\beta}' (X - X^*)) \quad (13)$$

2.5. Accelerated Failure Time Model (AFT)

According to [10], model describes the relationship between a set of covariates and the survival probabilities. For a group with predictor variables or covariates $X = (x_1, x_2, \dots, x_n)$, where t_i represents the time at which individual i experiences a particular event or right censoring ($i = 1, 2, \dots, n$). the AFT model is written mathematically as

$$S(t) = S_0(\phi t) \quad (14)$$

where $S_0(t)$ is the baseline survival function and ϕ is an acceleration factor that is, a ratio of survival times corresponding to any fixed value of $S(t)$. The acceleration factor is given according to the formula

$$\tilde{\phi} = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i \quad (15)$$

Under AFT model, the covariate effects are assumed to be multiplicative on the time scale and the covariate impacts on survival by a constant factor (acceleration factor).

The AFT regression perspective can also be formulated with respect to the random variable T , rather than to $\log T$, for convenience of specifying a likelihood function.

$$T = \exp(\mu + X\beta) \exp(\phi\epsilon) \quad (16)$$

$$\ln T = \mu + X\beta + \epsilon \quad (17)$$

Suppose that $Y = \log T$ is linearly associated with the covariate vector x . Then

$$Y = \mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_i x_{ri} + \tilde{\phi} \epsilon_i \quad (18)$$

$$Y - x_i \beta_i = \mu + \tilde{\phi} \epsilon_i \quad (19)$$

where μ is intercept, β 's are the estimates of the covariates x , ϕ is scale parameter and, ϵ_i is a random variable, T is the random variable assumed to have a particular distribution with survival function S , cumulative density function (cdf) F , and probability distribution function (pdf) f . where $\exp(\phi\epsilon) > 0$ has the hazard function $h_0(\tilde{e})$ and is independent of β . Given the equation $T \geq t$, the survival function for the i th individual can be mathematically expressed by

$$S_i(t) = P(T \geq t) = P(\exp(\mu + x' \beta + \tilde{\phi} \epsilon_i) \geq t) \quad (20)$$

the term $x' \beta$ is independent of the disturbance parameter, the survival function $S(t)$ can be written by two separate components;

$$S(t) = S_0(t \exp(x' \beta)) \quad (21)$$

Now,

$$S(t) = \Pr(T > t | x) \quad (22)$$

$$S(t) = \Pr(Y > \ln t | x)$$

$$S(t) = \Pr(Y - x' \beta > \ln t - x' \beta | x) \quad (23)$$

$$S(t) = \Pr(\exp(Y - x' \beta) > \exp(\ln t - x' \beta) | x)$$

$$S(t) = \Pr(\exp(\mu + \varepsilon) > t \exp(-x' \beta) | x) \quad (24)$$

Comparing equation 21 and 24. We have

$$S(t | x, \beta) = S_0(t \exp(-x' \beta)) \quad (25)$$

According to the relationship of survival function as $S(t) = \exp(-\beta x)$ and the hazard function can be expressed in terms of probability distribution function (pdf) and survival function $S(t)$ as follows

$$h(t | x, \beta) = \left[\frac{f(t | x, \beta)}{S(t | x, \beta)} \right] \quad (26)$$

$$h(t | x, \beta) = \frac{\exp(-x' \beta) f_0(t \exp(-x' \beta))}{S_0(t \exp(-x' \beta))} \quad (27)$$

$$h(t | x, \beta) = \exp(-x' \beta) h_0(t \exp(-x' \beta)) \quad (28)$$

Let $g(x | \beta) \exp(-x' \beta)$ and $Z(t) = t \exp(-x' \beta)$, thus

$$h(t | x, \beta) = g(x | \beta) h_0(Z(t)) \quad (29)$$

$$f(t | x, \beta) = g(x | \beta) f_0(Z(t)) \quad (30)$$

$$S(t | x, \beta) = S_0(Z(t)) \quad (31)$$

2.5.1. Generalised Weibull-Lindely (GWL) Model

Let the random variable X follow Lindley distribution with parameter β , the probability distribution function (pdf) and cumulative density function (cdf) of Generalized Weibull-Lindley (GWL) Distribution are given as follows:

$$f(x) = \left\{ \exp \left\{ - \left\{ \frac{-\log \{ (1 - ((\theta + 1 + \theta x) / (\theta + 1)) \exp(-\theta x)) \}}{\beta} \right\}^k \right\} \right\} \quad (32)$$

For $x \geq 0, (\theta, \beta, k) > 0$

$$F(x) = 1 - \exp \left\{ - \left\{ \frac{\log \{ (1 - ((\theta + 1 + \theta x) / (\theta + 1)) \exp(-\theta x)) \}}{\beta} \right\}^k \right\} \quad (33)$$

Survival function for GWL distribution is given as

$$\begin{aligned} S(x) &= 1 - F(x) \\ &= \exp \left\{ - \left\{ \frac{\log \{ (1 - ((\theta + 1 + \theta x) / (\theta + 1)) \exp(-\theta x)) \}}{\beta} \right\}^k \right\} \end{aligned} \quad (34)$$

In terms of the survival model, the GWL hazard function can be formulated as an inverse given by

$$\begin{aligned} H(x) &= \ln S(x) \\ &= -\ln \left(\exp \left\{ - \left\{ \frac{\log \{ (1 - ((\theta + 1 + \theta x) / (\theta + 1)) \exp(-\theta x)) \}}{\beta} \right\}^k \right\} \right) \end{aligned} \quad (35)$$

From the equation (35), we have

$$H(x) = - \left\{ \frac{\log \{ (1 - ((\theta + 1 + \theta x) / (\theta + 1)) \exp(-\theta x)) \}}{\beta} \right\}^k \quad (36)$$

We will consider Weibull model, Log-logistic model, Generalized Weibull-Lindely and Exponential model. For the purpose of comparison, Akaike Information Criterion (AIC) will be used to determine the most appropriate model for the data. The AIC [1], is given by

$$AIC = -2ll + 2(p + k) \quad (37)$$

Where ll is the log-likelihood (the probability of the data in a given model), p is the number of parameters/covariates, $k = 1$ for exponential model and $k = 2$ for Weibull and Log-logistic model [11]. Smaller AIC indicates better likelihood.

3. Results

The study aimed at identifying factors that are associated with and are predictive of survival of neonatal jaundice patients. Thus, this chapter contains three sections; the section one describes variables of interest and applies of a non-parametric approach to explain the nature of the variables and how they relate to the survival of neonatal jaundice patients. The second section presents analysis on the PH assumption test for variables and the models. The third section presents the models selection process and the model used for interpretation in this study.

Table 1. Variable Description

Variable	Description	Categories
Gender	Sex of patients	Male/Female
Types of Delivery	Mode of delivery of babies by Mother	VD=1,SVD=2,CS
Use of Traditional Medicine	Usage of herb Medicine during pregnancy	No=0 Yes=1
Antenatal Care	Usage of health care facility by Mother	No=0 Yes=1
Mother's Blood Group	Blood Group of the Mother	"A","B","AB","O"
Child's Blood Group	Child's blood Group	"A","B","AB","O"
Mothers' Age	Patients Mothers' Age	Numeric
Weight	Patients' Weight	2.5, ≥ 2.5

The study used a non-parametric approach to provide a summary of the distribution of the variables. Table 1. Presents description variables and categorization of the covariates for clearer understanding. From table 2, the time to surviving neonatal Jaundice overall median is found to be 45 days, which indicated that 50% of the neonatal patients survived jaundice in less than or equal to 45 days and the other 50% survived jaundice longer than 45 days. This is the survival time at which the cumulative survival function is equal to 0.5. This is summarized in Figure 1.

Table 2. Demographic and Medical profile of Mother and Patients (Neonates)

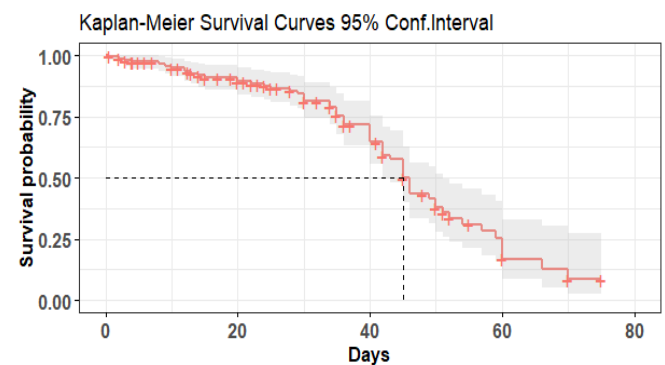
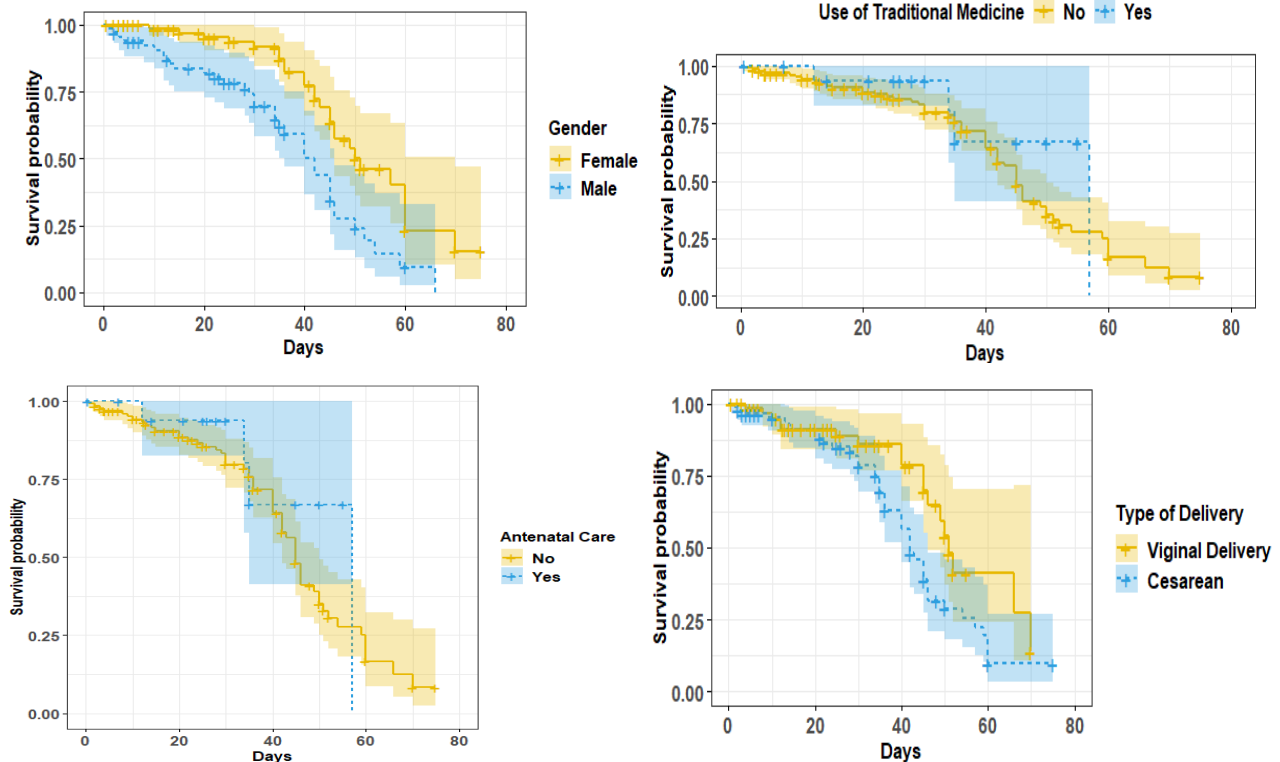
		Freq.	%	Mean \pm SD
Sex	Female	86	57%	
	Male	54	43%	
Use of Tradition Medication	No	132	88%	
	Yes	18	12%	
ANC Visit	No	69	46%	
	Yes	81	54%	
Weight	< 2.5	56	37%	
	+ 2.5	94	63%	
Type Of Delivery	VD	65	43%	
	CS	85	56%	
Mother Blood Group	A	28	19%	
	B	32	21%	
	AB	13	9%	
	O	77	51%	
Child Blood Group	A	56	37%	
	B	43	29%	
	AB	5	3%	
	O	46	31%	
Mothers' Age		29 \pm 6.24years		

The minimum and maximum survival times for neonatal jaundice patients were 0 to 80 days. Eighty days is the predetermined period for which patients who exceeded 80 days after diagnosis are termed to have survived. Figure 1

depicts the Kaplan-Meier probability of the survival of the Neonatal Jaundice patients with a 95% confidence bound, thus the plot observed that the survival probability of the disease decelerate from days 5 above so it is advice that the mother has to be at alert and monitor the child carefully between 0 to 5day of their lives.

Table 3. Cox Regression and parametric Models with prognostic factors

Model	Log-Likelihood	AIC
Exponential AFT	-112.0500	242.1003
Weibull AFT	-90.86353	201.7271
Log-Logistic AFT	-97.49812	214.9962
Generalized Weibull AFT	-76.49116	180.9830
Cox PH	-223.3567	470.7534

**Figure 1.** Survival function of days to survival after the diagnosis of the disease

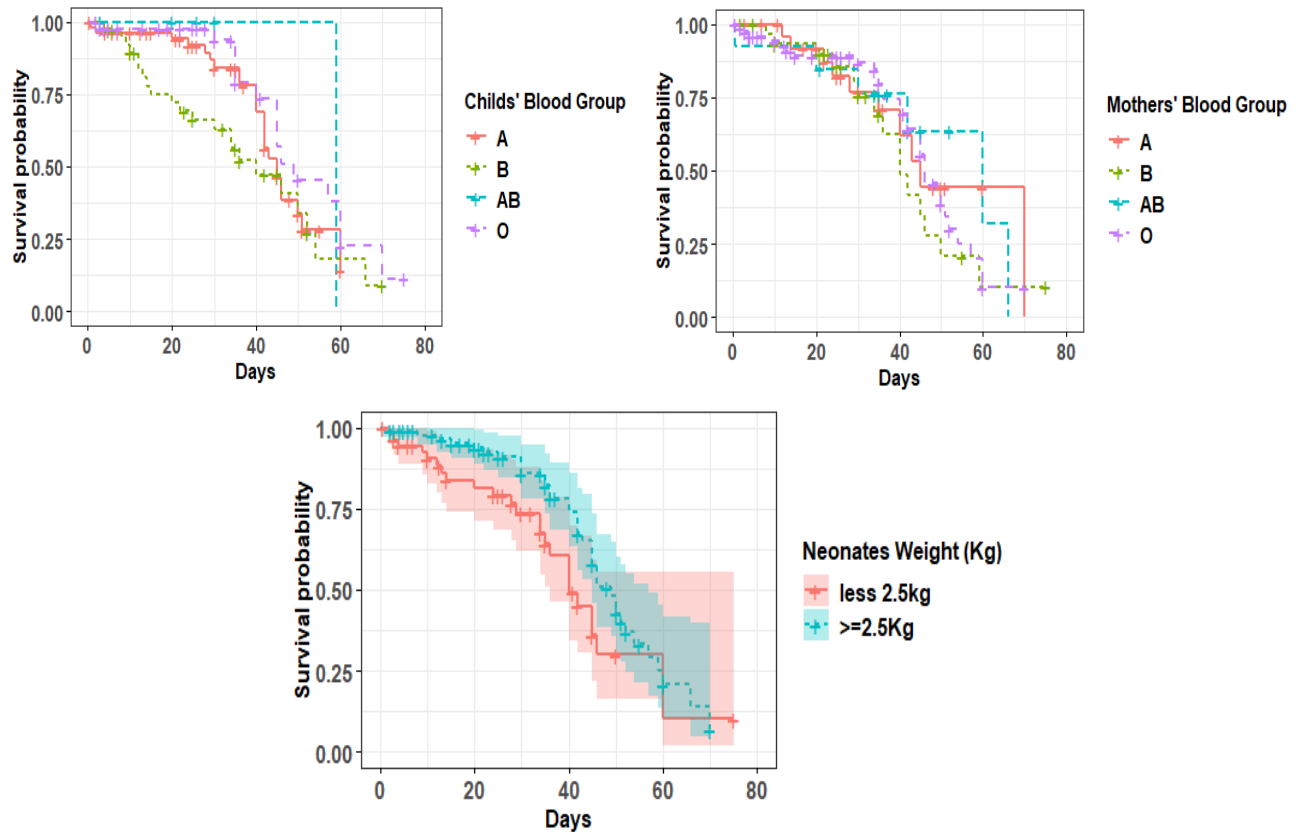


Figure 2. Survival function of time to survival by covariates

Table 4. Generalized Weibull Lindley Model

Covariates	Coef.	Std.Err.	Z	P> z	[95% Conf.Interval]	
UseofTrad_Medication (Yes)	0.20335	0.304476	0.67	0.504	-0.39341	0.800113
Antenatal Care Visit (No)	-0.50052	0.191061	-2.62	0.009	-0.87499	-0.12605
Mothers' age	-0.00281	0.014842	-0.19	0.85	-0.0319	0.026276
Weight (≥ 2.5 kg)	0.405128	0.203097	1.99	0.046	0.007066	0.80319
Type Of Delivery	-0.58344	0.194076	-3.01	0.003	-0.96382	-0.20306
Mothers' Blood Group						
B	-0.21797	0.291894	-0.75	0.455	-0.79008	0.35413
AB	-0.47093	0.353333	-1.33	0.183	-1.16345	0.221589
O	-0.45672	0.257191	-1.78	0.076	-0.9608	0.047368
Childs' Blood Group						
B	-0.40467	0.208406	-1.94	0.052	-0.81313	0.003803
AB	0.842851	0.7618	1.11	0.269	-0.65025	2.335952
O	0.277592	0.229192	1.21	0.226	-0.17162	0.726801
Gender (Yes)	-0.6663	0.203102	-3.28	0.001	-1.06437	-0.26823
_cons	4.991072	0.527752	9.46	0	3.956698	6.025447

4. Discussion

The aim of this research is to identify significant factors that are associated with and are predictor of survival of neonatal jaundice patients. The data used for this study was sourced from Ahmadu Bello University Teaching Hospital, Shika, Zaria. This study used a secondary data collected from the clinic records of 150 Neonatal jaundice patients.

The study utilized history of patients or medical records from year 2019 to year 2021. Clinical data such as presentation of the patient, use of traditional medication, birth weight, mother's blood group, child's blood group, mode of delivery, and Antenatal Visit Clinic (ANC) visit were collected. Kaplan Meier (K-M) was used to estimate the survival function of neonatal jaundice patients and determine the significant factors associated with the neonatal jaundice

disease using the selected parametric distribution.

Kaplan Meier (KM) curves are plotted for seven of the categorical variables with time recorded in days. The categorical variables of the study includes Gender, Use of traditional medicine, Use of Antenatal Care, Type of Delivery, Neonates' Blood group, Mothers' Blood group and Weight. From table 2, the Kaplan Meier plot for gender, indicated that the female neonates have higher survival probability than male neonates. But with a cautious observation, the curve for the two gender are very close. Though, in the earlier findings, it was revealed that the chances of female neonates have shorter median time of surviving compared to male neonates [6].

From Figure 2 it can be observed that patients who used traditional medicine have a higher probability of surviving than patients who did not employ traditional medicine the difference in survival for the two groups runs through the study period. Patients who did not visit Antenatal care have shorter survival time than those who used the Antenatal care. From the plot in Figure 2, the difference in the survival plots for the two groups are quite wide.

By looking at the child's blood group in Figure 2, the plot shows that neonates with A blood group has a higher survival probability compared to those with B, AB and O blood group. The difference in survival is visible until after the 70th day where the difference became visible. Considering the neonates' mothers' blood group, those with B and AB blood group have higher probability of surviving than other blood group. The survival plots for the groups are quite close suggesting that the differences are not significant enough. This implies that differences between these curves are significant enough. In Figure 2, the plot shows that patients with over 2.5Kg weight have a shorter survival probability than those weighing less than 2.5Kg.

In comparing with the Parametric Accelerated Failure Time (AFT) models which are Exponential, Weibull and Log-logistic models, as against the Generalized Weibull Lindley and Cox proportional Hazard model. The models will be compared on the basis of their AIC values to select the best-fitted model and finally the selected model will be used for interpretation. Table 3 presents the hazard ratios for the Semi parametric model (Cox PH), and also the Time Ratio (TR) for the parametric models (Exponential, Weibull and log-logistics and GWL). Generally, AIC is used to select the best models. The lowest AIC leads us to identify the best one. According to the result, all the parametric models had the lowest AIC indicating that they outperform in comparison to Cox PH regression model, and among the parametric models, the Generalized Weibull Lindley (GWL) outperformed the rest of the parametric models by having a lowest AIC of 180.92 as seen on Table 3. Thus, GWL model was considered the best model that fits the data. However, in the studies of [6], [11] GLW appeared to be more flexible when compared with other distribution.

The final results for the GWL AFT model was shown in Table 4 and the result revealed that the survival time of neonates jaundice patients is statistically significantly

influenced by Antenatal visit, weight, type of delivery and gender respectively.

5. Conclusions

The Generalized Lindley-weibull (GLW) was studied and applied to neonatal jaundice data. The result of the application based on the performance measures, showed that the Generalized Lindley-weibull (GLW) model is the best fit for modelling neonatal jaundice data. Kaplan Meier (KM) plots for the seven categorical variables with time recorded in days indicated that patients who used traditional medicine will have a higher probability of survival time than patients who did not employ traditional medicine and patients who did not visit Antenatal care will have a shorter survival time than those who used the Antenatal care. The results for the GWL AFT model showed that the survival time of neonates jaundice was statistically significantly influenced by Antenatal visit, weight, type of delivery and gender respectively.

REFERENCES

- [1] Akaike H. (1973). "Information Theory and Extension of the Maximum Likelihood Principle ". In B.N. Petrov and F. Csaki (eds) 2nd *International Symposium on Information Theory*, Academia Kiado, Budapest 267 – 281.
- [2] Alzaghal A., Lee C., and Famoye F. (2013). Exponentiated T-X family of distributions with some application. *International Journal Probability and statistics*. 2, 31-49.
- [3] Cox D. (1972). Regression Model and Life-table. *Journal of the Royal Statistical Society*, 34, 187-220.
- [4] Cox D. R. and Oakes D. (1984). *Analysis of Survival Data*. Chapman and Hall Limited, London.
- [5] Everitt B.S. (2002). *The Cambridge Dictionary of Statistics*. Second Edition. The Edinburgh Building, Cambridge, United Kingdom.
- [6] Folorunso S. A. and Osanyintupin O. D. (2018) Comparism of Cox Proportional Hazard model and Accelerated Failure Time Models: An Application to Neonatal Jaundice. *International Journal of Public Health and Safety*. 3(4) 100-171.
- [7] Grover, G., Sreenivas, V., Sudeep K. and Divya S. (2013). Estimation of survival of Liver Cirrhosis Patients, in the presence of prognostic Factors Using Accelerated Failure Time Model as an Alternative to Proportional Hazard Model. *International Journal of Statistics and Applications*, 3(4), 113-122.
- [8] Johansen S. (1978). The product limit estimate as a maximum likelihood estimate. *Scandinavian Journal of Statistics* 5, 195-199.
- [9] Joseph G.C. (2010) Survival analysis: Overview of parametric, nonparametric and semi-parametric approaches and new developments. SAS Global forum, Michigan State University, USA.

- [10] Kalbfleisch J.D. and Prentice R.L. (2002). *The Statistical Analysis of Failure Time Data*. Second Edition. Wiley, New York.
- [11] Klein JP, Moeschberger L.M. (2005). *Survival Analysis: Techniques for Censored and Truncated Data*. Second edition. New York.
- [12] Pius M. and Gadde S. R. (2020). Generalized Weibull-Lindley Distribution in modeling Lifetime Data. *Journal of Mathematics*, 1, 23-38.
- [13] Xian Liu (2012). *Survival Analysis models and Applications*. First Edition, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, United Kingdom.