Scientia Africana, Vol. 11 (No.2), December 2012. Pp 22- 29
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

APPLICATION OF SURVIVAL ANALYSIS ON MALARIA AMONG CHILDREN: A CASE STUDY OF UNIVERSITY OF ILORIN TEACHING HOSPITAL

J. I .Ariyo[1], A. A. Abiodun[2] and B. A. Oyejola[3]
Communication and Market Research Group Ltd.,
21 Ogunlowo Street,
Ikeja Lagos State, Nigeria[1]

Department of Statistics,
University of Ilorin,
P.M.B. 1515 Ilorin, Nigeria[2,3]
ariyo2006@yahoo.com[1], manfredwole@yahoo.com[2] and boyejola2003@yahoo.com[3]

Received: 18-10-12
Accepted: 20-11-12

ABSTRACT
This study examined the influence of some covariates age, sex, and presence of consultant on the risk of death from malaria among children, aged 15 years or less, admitted at University of Ilorin Teaching Hospital, Kwara State Nigeria. Cox proportional hazards model was used to analyze the data. Results from analyses showed that when age of the patients and presence of a consultant at the stage of admission were adjusted for, male patients had lower risk of death than their female counterparts. Those that met with consultants at admission stage had lower risk of death than those who did not. The risk of death for younger children aged 5 years or less was also found to be higher than for those older than 5 years. It was also observed that there were significant interaction effects between age and sex of the children as well as between age and presence of consultant at admission stage.

Key words: Survival Function, Hazard function, Kaplan Meier, Censoring, Partial likelihood.

INTRODUCTION
Survival time or failure time measures the length of time from a defined starting point to the occurrence of a given event. The event is often termed "failure event", where the word failure in survival analysis context is a generic term which does not necessarily mean failure in the literal sense. Examples of failure time include: time to acceptance of a job offer, time to termination of a mortgage loan, time to return to a criminal act by an individual who has been convicted of a prior offence, sentenced and corrected/released and time to a transition from single life to married life. In medical science failure time could be time from the beginning to the end of a remission period, time from the diagnosis of a disease to death, time to occurrence of or recovery from a disease. Standard statistical techniques cannot be applied in analyzing survival time data because the underlying distribution is rarely Normal. Survival analysis attempts to answer such questions as, what is the fraction of a population which will survive past a certain time. Of those that survive, at what rate will they fail? How does a particular circumstance or characteristic increase or decrease the odds of survival

To determine the survival time T, three basic elements are needed: a time origin or starting point, an ending event of interest and a measuring scale for the passage of time.

In many situations, we have additional information on individual subjects that we believe may be associated with risk of the event of interest occurring. Due to censoring, standard linear regression methods are not feasible in modeling the relationship between survival time and some biological, socio-economic and demographic characteristics that could possibly affect the survival status of patients (Kleinbaum and Klein, 2005). The survival time of an individual is said to be censored when the time of event of failure occurring to an individual is unknown. This may be because some individuals in the study are still alive at the time of data collection. Also, the survival date of an individual at the time of the analysis may be unknown because the individual has been lost to follow-up. For the censored individuals we know only that the time to failure is greater than the censoring time (Cox, 1972).

In survival time data the response (dependent) variable is often the time to an event of failure and in modeling survival data, the hazard function is often used. Since survival time is often skewed (non-normal), median is rather used as a measure of central location than the arithmetic mean which could be influenced by extreme values. One common procedure to estimate median survival time is by obtaining the Kaplan Meier curve. Then the median survival time can be obtained from it by starting at point 0.5 on the y-axis, projecting it horizontally until the survival curve is met, and heading down to the x-axis time point to obtain the estimated median survival time.

One major interest in analysis of survival time data is examining the effect of prognostic factors on the hazards function. Parametric model such as Weibull or semi parametric model such as Cox proportional model (Cox, 1972) are often used. A major constraint which determines the choice of method in modeling survival data is the shape of baseline hazards. When empirical information is sufficient, parametric models can provide some insight into the shape of the baseline hazard (Nardi and Schemper, 2003). Furthermore, extrapolations of survival functions become possible which – though speculative – may be of interest in some applications (Gelber et al.(1993).

However, often, either the true hazard is not known or it is complex, in which case assumptions underlying parametric models may not be true for such data. For example, when a Weibull model is used for analysis of data from a population that is not from a Weibull survival distribution, Weibull parametric model may produce estimates that are biased and inefficient. In such a situation, Cox model will be more efficient than the Weibull model parametric model analysis (Harrel, 2001)

MATERIALS AND METHODS

The distribution of the survival time T from a starting point to the event of interest is characterized by two functions of interest (Chap T. LE, 2003). These functions are Survival or Survivorship function and Hazard function. The survivorship function, conventionally denoted by S(t), is defined as the probability that a person survives longer than t units of time and is expressed as:

$$S(t) = \Pr(T > t), \qquad 2.1$$

where t is some time and *T* is a random variable denoting the time of failure. The hazard or risk

Scientia Africana, Vol. 11 (No.2), December 2012. Pp 22- 29
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

function gives the instantaneous failure rate and is defined as follows. Assuming that a typical patient has survived to time t, then for a small time increment dt, the probability of an event of failure occurring during time interval (t, t + dt) to that person is given by,

$$\lambda(t)\,dt = \Pr(t \le T < t + dt \mid T \ge t) = \frac{f(t)\,dt}{S(t)} = -\frac{S'(t)\,dt}{S(t)} \qquad 2.2$$

The hazard or risk function approximates the proportion of subjects dying or having events of interest per unit time around time t. It is also known as the force of mortality and is a measure of the proneness to failure as a function of the person's age (Nelson, 1982).

Cox Proportional Hazards Model

Cox model is a semi parametric model that makes no assumption regarding the nature of the baseline hazard function. However, its functional form for the dependency of the survival time on the covariates is fully parametric. Suppose $\lambda(t/X)$ is the hazard function at time t of an individual with covariate vector $X=[x_1, x_2, \ldots\ldots x_k]$, then the Cox proportional hazards model (Cox, 1972) is given by

$$\lambda(t / X) = \lambda_0(t)\exp(\beta' X) \qquad 2.3$$

where $\lambda_0(t)$ is an unspecified baseline hazard when all explanatory variables are zero, and $\beta=[\beta_1,\beta_2,\ldots,\beta_k]'$ is a vector of unknown regression coefficients.

Review of literature shows the extensive use of the Cox proportional hazards regression model for hazard rate or instantaneous risk of a given event (Orbe et al., 2002; Moran et al., 2008). The basic assumption underlying this model is proportionality of hazard rates (Sayehmiri,et al. 2008). The assumption states that the ratio of the hazards for two individuals is expected to be constant over time. However, where Proportional Hazards (PH) assumption is not met, it is improper to use standard Cox PH model as it may entail serious bias and loss of power when estimating or making inference about the effect of a given prognostic factor (Moran et al. 2008; Kleinbaum and Klein, 2005).

The value of $\exp(\beta_j)$ is called the hazard ratio or relative risk. It represents the risk associated with an exposure if x_j is binary (expose $x_j = 1$ versus unexposed $x_j = 0$) or the risk due to a 1- unit increase if x_j is continuous.

Model Comparison

Often, need arises to compare models. Several methods are commonly used for comparing models in Generalized Linear Models (GLM). One popular method tool is Akaike's Information Criterion (AIC), which is expressed as

$$AIC = -2\log(L+2p)$$

where L is the likelihood and p is the number of elements in the parameter vector. The model with smaller AIC is better.

Data Analysis

The data used in this study were collected from the Pediatric Department, University of Ilorin Teaching Hospital for the year 2009. Eight hundred and seventy seven in-patients malaria children whose ages were not more than 15 years, admitted at the hospital were covered by the study based on the daily register obtained from the hospital. Information was collected on age, sex, consultant (presence/absence of consultant at admission stage). Of these patients, 497(56.67%) were male while 380(43.33%) were female. Also 534(60.9%) were aged 5 years or less while 343(39.1%) were above 5 years of age. Information was also collected on the duration (in days) of hospital admission before death, which

defines the survival time and the death status (whether or not the patient had died). Patients who were still alive at the time of data collection and those that were transferred were right censored, because their death status could not be ascertained at the time of data collection.

The main aim of the study is to investigate the relationship between days of hospital admission before death of the malaria patients and the explanatory variables which are thought to be the influential factors.

In examining the effect of age and sex of patients and whether or not they met with consultant at admission stage, Cox proportional hazards model was used. At the first stage, models were fitted by incorporating each of the three covariates: age, gender and consultant in separate models. It was observed that none of them was found to be significant (results not shown). The model termed Model1was then fitted which incorporated the three covariates gender, consultant and continuous age in the same model. It was discovered that only age was significant in this model and was then dichotomized using the mean age (5 years) as threshold. Two models (Models 2 and Model 3) were then fitted using the dichotomized age along with covariates gender and consultant. While Model 2 incorporated the covariates as the main effects only, Model 3 also included the interaction effects among these variables. The results of the three models are presented in Table 2. These models were also compared using Akaike Information Criterion (AIC). For the study, gender was coded 1 for male and 0 for female, consultant was coded 1 if there was presence of consultant at admission stage and 0 if not, dichotomized age was coded 1 for a patient whose age was not more than 5 years and coded 0 for a patient whose age was more than 5 years. All analyses were done using STATA 10.

RESULTS

Results of Descriptive Analysis

The median duration time of admission before death was estimated using Kaplan Meier survival curves. Figures 1 and 2 show curves for overall patients and also separately for male and female patients. It is observed that the estimated median survival times for males and females are the same, which equals the overall median survival time (11 days).

Table 1 presents the overall mean age of the patients on admission and also the mean age by sex and by consultant. It also shows the results of test of equality of mean age between male and female patients as well as between those who met with consultant and those who did not at admission stage. From the table, the mean age of all the children is 5.2485±0.1772. The mean ages of male and female patients are 5.5092±0.1537 and 4.5022±0.2394 respectively and are significantly different from each other at 5% level of significance. It is also observed that the difference between the mean age of those who met with a consultant and those who did not at admission stage is not significant.

Results of Cox Proportional Model

From Model 1, none of age, sex or consultant has significant effect on the risk of death from malaria disease. In model 2 with age dichotomized, children aged 5years or less have higher risk of death than the older children when sex of patients and presence of a consultant at admission stage are adjusted for. Based on the values of AIC, Model 3 which included the interaction terms is the best model (AIC=1531.044) and further discussions will be based on this model. As observed, the effects of age, sex and consultant are observed to be significant when interaction terms are included. Adjusting for age and presence of a consultant, male patients have lower risk of death than their female counterparts (Haz.

Scientia Africana, Vol. 11 (No.2), December 2012. Pp 22- 29
 ISSN 1118 – 1931

Ratio= 0.7701, P-value=0.014). Those that met with consultants at admission stage have lower risk of death than those who did not (Haz. Ratio= 0.8600, P-value=0.044). Age is also observed to be significantly associated with the risk of death under this model (with interaction terms). The risk of death for younger children aged 5 years or less is 1.4 times higher than those older than 5 years. It is also observed that there are significant interaction effects between age and sex of patients as well as between age and consultant.

Fig. 1. Overall Kaplan Meier survival curve

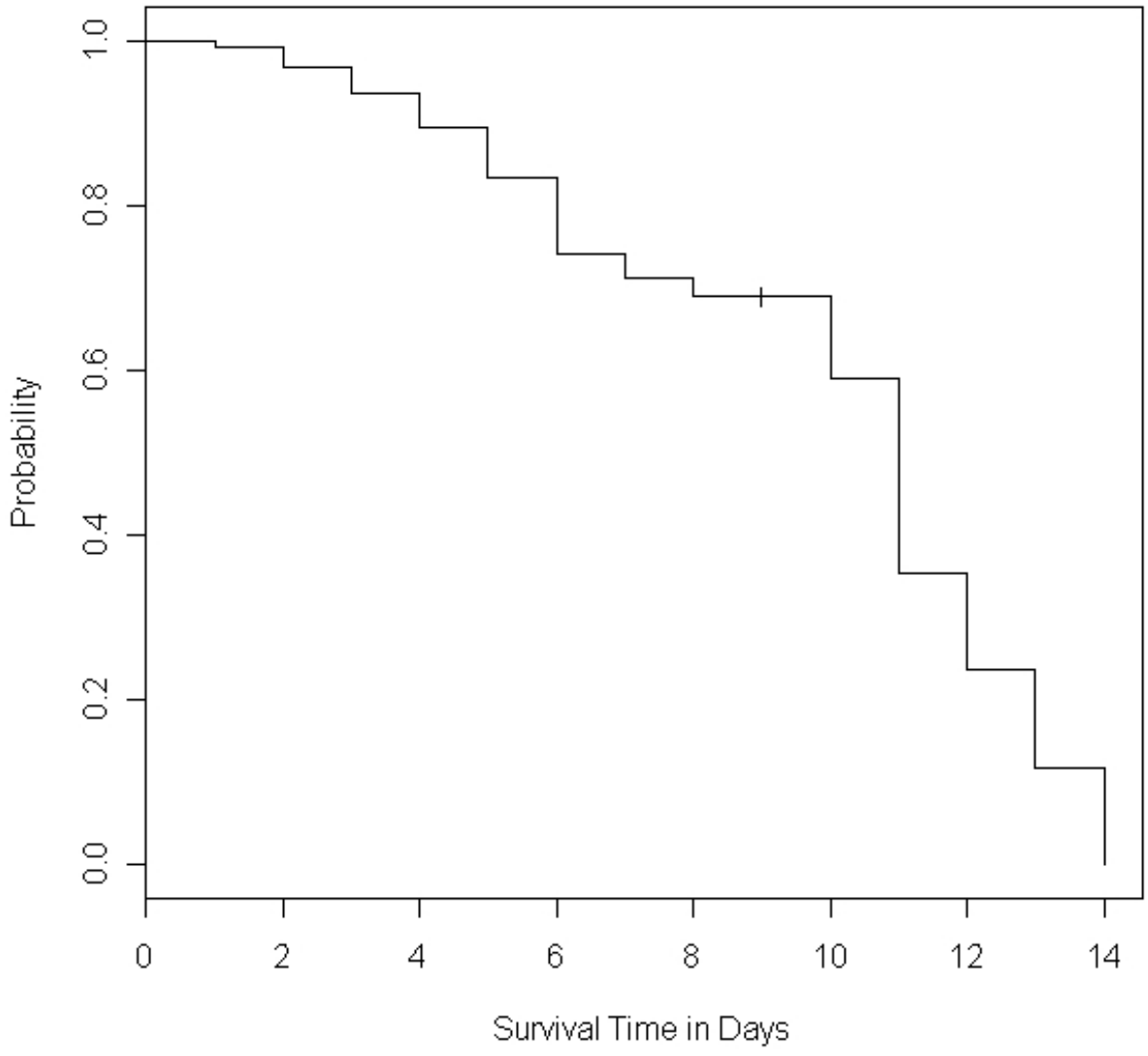

Fig. 2. Kaplan Meier survival curve by sex

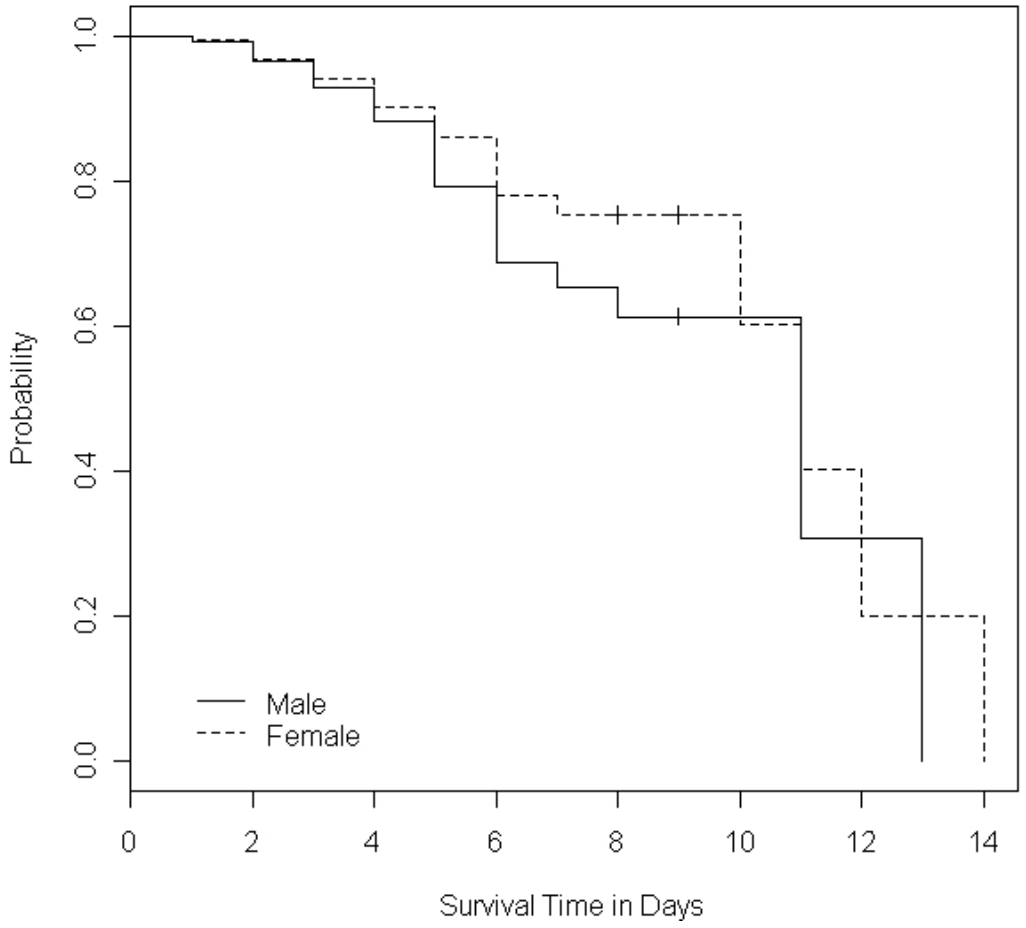

Table 1: Overall mean age and mean age by sex and by consultant.

Variable	Mean±se(mean)	T	$Pr(\|T\|) > (\|t\|)$
Overall	5.2485±0.1772		
Sex			
Male	5.5092±0.1537	3.4015	0.001
Female	4.5022±0.2394		
Consultant			
Present	5.0262±0.1645	-1.9536	0.108
Absent	5.5395±0.2099		

Scientia Africana, Vol. 11 (No.2), December 2012. Pp 22- 29
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

Table 2: Hazard Ratios, P-values and AIC for the three models

Covariates	Model1		Mode2		Mode3	
	Haz. Ratio	P-value	Haz,. Ratio	P-value	Haz. Ratio	P-value
Age	0.7144	0.420	-	-	-	-
Male	0.8933	0.124	0.8939	0.123	0.7701	0.014
Consult	0.8357	0.365	0.8368	0.358	0.8600	0.044
5years	-	-	1.2168	0.041	1.3781	0.032
5years*Male	-	-			0.6148	0.011
5years*Consult					0.8905	0.018
Male*Consult					0.8402	0.659
AIC	1550.206		1539.321		1531.044	

DISCUSSION

The study analyzed data collected on Malaria children whose ages were not more than fifteen years, from University of Ilorin Teaching Hospital. Information was collected on duration of admission before death, which defined the survival time. Information was also collected on three key explanatory variables that were thought to be capable of influencing the survival experiences of these children. These include age of the children, sex and whether or not there was presence or absence of a Consultant at admission stage.

Kaplan-Meier (KM) survival curve was used to estimate the median survival time.

Cox Proportional Hazards Model was used to examine the relationship between the risk of death from malaria and the explanatory variables collected. From the descriptive analysis, it was found that the median survival time for male patients was similar to that of female patients which was also similar to overall median survival time for all patients.

Based on the values of AIC, model that incorporated the interaction terms among the explanatory variables used in this study was the best in modeling the survival experience of the malaria children. Finding from analyses showed that when age of the patients and presence of a consultant at the stage of admission were adjusted for, male patients had lower risk of death than their female counterparts. It was also discovered that those that met with consultants at admission stage had lower risk of death than those who did not. The risk of death for younger children aged 5 years or less was also found to be higher than those older than 5 years. It is also observed that there were significant interaction effects between age and sex of patients as well as between age and consultant.

REFERENCES

Chap T. LE (2003). Introductory Biostatistics. Cancer Center University of Minnesota, 380-500.

Cox D.R (1972). Regression models and Life-tables (with discussion). *J. Roy. Statist. Soc. Ser. B. 34, 187-220.*

Gelber R.D. Goldhirsch, A. and Cole, B.F.(1993). Parametric extrapolation of survival estimates with applications to quality of life evaluation of treatments. *Controlled Clinical Trials* ; 14:485-499.

Harrel, E.H.(2001). Regression Modelling strategies (With Applications to Linear Models, Logistic Regression and Survival Analysis). *Springer-Verlag NewYork*, Inc.

Kleinbaum, D.G. and Klein, M. (2005). Survival Analysis – A Self-Learning Text. Springer-Verlag.

Moran, J.L. Bersten, A.D. Solomon, P.J. Edibam, C. and Hunt, T.(2008). Modelling survival in acute severe illness: Cox versus accelerated failure time models . *J Eval Cin Pract*. 14:83–93.

Nardi, A. and Schemper, M. (2003). Comparing Cox and Parametric Models in Clinical Studies. *Statistics in Medicine* 22, 2597-3610.

Nelson, W. (1982) Applied Life Data Analysis: 23-24. John Wiley & Sons, Inc. New York.

Orbe, J. Ferreira, E. and Nunez-Anton, V. (2002). Comparing proportional hazards and accelerated failure time models for survival analysis. *Stat Med*. 2002;21:34933510. doi: 10.1002/sim.1251.

Sayehmiri, K. Eshraghian, M. R. Mohammad, K. Alimoghaddam, K. Foroushani, A. R. Zeraati, H. Golestan, B and Ghavamzadeh, A.(2008). *J. Exp. Clin. Cancer Res. 27(1):27-74.*