A SURVIVAL ANALYSIS ON THE INCIDENCE OF PNEUMONIA AND INFANT MORTALITY RATE IN EDO CENTRAL SENATORIAL DISTRICT OF NIGERIA

Izomo E. J.[1*], Ikpotokin O.[2] and Ogbeide E. M.[2]

[1]Department of Statistics, School of Applied Sciences, Edo State Polytechnic Usen, Nigeria.
[2]Department of Mathematics and Statistics, Faculty of Physical Sciences, Ambrose Alli University, Ekpoma, Nigeria
*Corresponding author. E-mail: julius_izomo@yahoo.com.

Accepted 8th June, 2020

The aim of this work is to study the epidemiological aspects of pneumonia and identify the risk factors for incidence rate of pneumonia among children aged 0-5 years treated as in-patients, at Irrua Specialist Teaching Hospital, Edo State, Nigeria. Adopting survival analysis technique, the number of patients that survived or remained after intervention over a period of time with the Kaplan-Meier estimator was computed and curves of survival probabilities were plotted. Statistical difference among the sex of the in-patients were compared and checked using the Log-rank test. Results obtained from the survival distribution reveal that as the number of days (time) increases, the chances of survival decline rapidly with a p – value of 0.03. The study shows that at 70.8 days, the chance of survival was 0.95 but as the days increased to 197.1 the chances of survival declined rapidly to 0.5. Mortality trend from the result also indicates that survival probabilities declined with the passage of time, and that the mean survival rate of male infants is higher than that of females.

Key words: Pneumonia, survival distribution function, mortality rate, Kaplan-Meier estimator, survival probabilities, log- rank test.

INTRODUCTION

Nigeria is one of the nations of the world with high pneumonia mortality in children (Udofia and Okonofua, 2008). Mortality refers to a death that occurs within a population at a given period of time. Infant mortality is the number of infants' death occurring within one year after birth per one thousand live births for a given year. Universally, childbirth event attracts celebration but due to the high risk in childbirth delivery encountered by both mother and child, the tragedy of deaths does occur especially during the first few days after childbirth (Ogbonaya and Aminu, 2009). Annually, Nigeria records one quarter of babies born not surviving within their first 28-30 days and before their fifth birthday, Ogunjimi et al. (2012). This unhealthy trend has become a matter of great concern, calling for concerted approach from all and sundry. Pneumonia, against the natural belief, is not caused by exposure to cold weather,but rather by a common bacterium known as *Streptococcus pneumonia* (Okumale, 2017). Pneumonia is a leading cause of morbidity and mortality in children 0-5 years old; it is responsible for approximately ratio 1:5 mortality cases in Nigeria (Kuti and Oyelami, 2015).

An increase in mortality rate shows a decline in the health status of a population and this is largely attributed to non- availability of health care; while a reduction in mortality signifies a tremendous increase in health care delivery services (Adetoro and Amoo, 2014). This is a major challenge in the Nigerian health care sector, hence a number of preventable diseases lead to death most especially in infants**.** Child mortality is associated with categories of acquired ailments of infectious diseases of which pneumonia has claimed the lives of many before their fifth birthday (Katrona and Katona, 2008). This poses a great danger for our children, but with global immunization advocacy this infectious disease can be prevented. The survival of our children basically is dependent on adequate health care facilities, and the absence of this factor poses a health risk and hazard to infants who are vulnerable to this disease (Kuti and Oyelami, 2015).

Children living in rural settings stand greater risk of not surviving than those in urban settings due to shortfalls in health programmes. If not given the needed attention by government policies on health care, it will lead to a drastic increase in morbidity and mortality (Finlay et al. (2011). Okumale (2017) advocates that unclean and unhygienic environments are breeding grounds for germs and bacterial infections which easily contaminate children. In Nigeria, most poor people live and give birth in unsafe and unclean environments making their babies more vulnerable to childhood killer diseases contacted from germs and bacterial infections. Pneumonia usually starts when germs are breathed into the lungs. Bacteria get into the body either through the mouth or other openings in the body and they contaminate the blood and respiratory apparatus. This makes it hard for the lungs to fight the infection. The likelihood of contacting the disease comes after having a cold or flu. The signs and symptoms include fever, chills, cough, shortness of breath and fatigue. These symptoms are followed by coughing out mucus sputum which is rusty, greenish or tingled with blood, sharp chest pain, shaking, teeth chattering, increased respiratory rate, nausea, vomiting, weakness of the body and diarrhea (Okumale, 2017).

Based on these challenges, this research introduces a statistical approach to the incidence of pneumonia through the use of survival analysis to ascertain the mortality rate. The specific objectives of this work are to determine the incidence rate of pneumonia in infants, its occurrence and impact, whether it is on the increase or decrease over the given time period and to investigate long term survivals of patients admitted to a public hospital with diagnosis of pneumonia in Edo Central Senatorial District of Edo State, Nigeria. The Kaplan-Meier together with the log rank test was adopted in this study because of its assumption free property in survival analysis. It is used to estimate conditional probabilities at each time an event occurs and take the product limit of the probabilities to estimate the survival rate at each point in time.

THE KAPLAN-MEIER METHOD AND THE LOG RANK TEST STATISTIC

Survival analysis is commonly used in clinical trials and biomedical sciences as a statistical tool in which the response variable is time. It is generally defined as a set of methods for analyzing data where the outcome variable is the time taken for the occurrence of a disease or death (Simona, 2008; Kleinbaum and Klein, 2012). The dependent or response variable is the waiting time until the occurrence of a well-defined event of death (Dafni, 2011). Statistical analysis of survival data shows the magnitude of the expected increase or decline in mortality from clinical trials (Ahmed et al., 2007). Two functions that are dependent on time in survival analysis are the survival function and the hazard function (Hosmer et al., 2008). The survival function gives for every time the probability of surviving that takes into account cases of survivorship; while the hazard function gives the potential that the event of death will occur per time unit an individual has survived up to the specified time (Hosmer et al., 2008).

The non-parametric estimator of the survival function known as the Kaplan - Meier method is used to estimate the proportion of surviving by any time. It is used to obtain univariate descriptive statistics for survival data (Baulies et al., 2015). The Kaplan-Meier estimator method of non-parametric statistics is also called a non-parametric maximum likelihood estimator used for estimating survival probabilities. The important assumption of the Kaplan-Meier survival function is that the distribution of censoring times is independent of the exact survival times and it accommodates no censoring. A data is said to be censored when values of the variable are not observed for some of the items in the sample. Patients may have censored survival time if death or recurrence has not yet occurred and this could happen when they drop out of the study or stop attending clinics for follow up. Similarly, certain individuals may drop out from the study or be lost to follow up. Each of these cases is said to be censored while non-censored data are cases where the data entry is complete and the patients complete the treatment.

The Kaplan-Meier estimator is assessed by measuring the number of subjects that survived after intervention over a period of time. The time starting from a defined point to the occurrence of a given event (death) is called the survival time and the analysis of group of data as survival analysis. The Kaplan-Meier method is a non-parametric estimator which involves computing of probabilities of occurrence of event at a certain point of time. It is widely used in clinical trials because of its versatility in estimating a population survival curve from a sample. In instances where every patient is followed until death, the curve may be estimated simply by computing the fraction surviving at each time. One unique feature of the Kaplan-Meier method is that it allows censoring and non-censoring. That means it allows estimation of survival over time even when patients drop out or are studied for different length of time. It works for each interval as survival probability is calculated by the number of patients surviving divided by the number of patients at risk or did not survive or dropped out. For large samples the Kaplan Meier method is approximately normally distributed with mean s(t) and variance $\hat{V}\left(\hat{S}(t)\right)$, the Kaplan - Meier estimator of the survivorship function or survival probability is given by:

$$s(t) = \rho(T > t) \quad (1)$$

is defined as

$$\hat{s}(t) = \prod\nolimits_{t_i \le t}\left(1 - \frac{d_i}{R_i}\right) \quad (2)$$

where ŝ(t) = survivorship function, $\prod_{t_i \le t}$ = product sum of ordered time, t_i= *ith* ordered follow-up time, d_i = number of deaths at *ith* ordered time, R_j = number of uncensored observation at *ith* ordered time , R_i = number of subjects at risk at *ith* ordered time and the Green wood's variance estimator is denoted by:

$$\hat{V}\left[\hat{S}(t)\right] = \hat{S}(t)^2 \sum_{t_{(i)} \le t} \frac{d_i}{R_i(R_i - d_i)} \quad (3)$$

In order to derive two approximate 95% confidence intervals for *s(t)* for a fixed *t*, or in general $(1-\alpha)100\%$, the lower and upper limits are denoted as follows:

$$LowerLimit = \hat{S}(t) - 1.96.\hat{S}(t)\sum\nolimits_{t_i \le t}\sqrt{\frac{d_i}{R_i(R_i - d_i)}} \quad (4)$$

and

$$UpperLimit = \hat{S}(t) + 1.96\hat{S}(t)\sum\nolimits_{t_i \le t}\sqrt{\frac{d_i}{R_i(R_i - d_i)}} \quad (5)$$

To compare the probability of survival beyond a certain time for two groups of subjects, the Z statistic is appropriate for the given test hypothesis:

$$H_o : S_1(t_o) = S_2(t_o) \quad (6)$$

$$H_1 : S_1(t_o) \neq S_2(t_o) \quad (7)$$

and

$$Z = \frac{\hat{S}_1(t_o) - \hat{S}_2(t_o)}{\sqrt{\hat{V}\left[\hat{S}_1(t_o)\right] + \hat{V}\left[\hat{S}_2(t_o)\right]}} \quad (8)$$

Where Z = Z test statistics, $\hat{S}_1(t_o)$ = survival function of group 1, $\hat{S}_2(t_o)$ = survival function of Group 2, $\hat{V}$ = variance of the survival function, H_o :there are no differences in survival distributions, H_1 : there are differences in survival distributions.

To test for overall differences between estimated survival of two or more groups of subjects, such as males versus females, or treated versus untreated, survivals or deaths, the log rank test comes to mind. The log rank test is a method used for comparing the Kaplan - Meier estimate for each group of subjects (George et al., 2014).

Log rank test statistic

It is a confirmatory test used to compare the entire survival function for two groups of subjects. It is more powerful because its analysis is based simply on proportions. The Log rank test is a comparing survival function used for each expected observed number of deaths in each group. It is used to compare the total expected death (ϵ_j) in each group to the total observed death (O_j). It is a type of chi square test used to assess the overall difference in survival analysis as denoted by:

$$\varepsilon_{ij} = \left(\frac{R_{1j}}{R_{1j} + R_{2j}} \right)(d_{ij} + d_{2j}) \qquad (9)$$

where for each j defined, ε_{ij} = expected number of deaths, d_{1j} = number of deaths in Group 1 d_{2j} = number of deaths in Group 2, R_{1j} = the number at risk in set Group 1, R_{2j} = the number at risk in set group 2. The log rank test statistic is denoted by:

$$\chi^2 = \sum_{i=1}^{2} \frac{(O_1 - E_1)^2}{\hat{v}_1} \qquad (10)$$

such that

$$\hat{V}_1 = \sum_{j=1}^{k} \frac{R_{1j} R_{2j} (d_{ij} + d_{2j})(R_{ij} + R_{2j}) - d_{ij} - d_{2j}}{(R_{ij} + R_{2j})^2 (R_{ij} + R_{2j} - 1)} \qquad (11)$$

and the test hypothesis is given by

$$H_o : S_1(t) = S_2(t) \text{ for all } t \qquad (12)$$

$$H_1 : S_1(t) \neq S_2(t) \text{ for some } t \qquad (13)$$

where χ^2 = chi square of observed deaths, $\hat{V}_1$ = variance of the number of deaths and the number at risk, H_o = there are no differences in the survival function of the two groups, and H_1 = there are differences in the survival function of the two groups.

Mean survival rate (MSR)

To measure the mean survival rate, the percentage of people in a study or treatment group still alive for a given period of time after diagnosis is taken into consideration. It is a statistic that describes how long an average person will survive for a particular amount of time. It is the total number of deaths (d_{ij}) in a defined time period divided by the total number of persons in the population (R_{ij}) at the beginning of the time period of the experiments or treatment group multiplied by 100%; it is denoted by:

$$MSR = \frac{d_{ij}}{R_{ij}} \times \frac{100}{1} \qquad (14)$$

METHODOLOGY

The methodologies used in this study are a non parametric estimator of the survival function known as the Kaplan-Meier survival function and the Log Rank test. The data for this paper are secondary data from Irrua Specialist Teaching Hospital, Irrua, Edo State, Nigeria. The data are on reported cases of children treated, survived and did not survive for 2007-2016 periods. In order to facilitate the computational efficiency, R-Statistical software is used in this paper to implement the method. The R-Statistical software has a package known as "*Survival"* used for survival analysis.

Data presentation and analysis

Data used in this study are obtained from Irrua Specialist Teaching Hospital (ISTH), Irrua, Edo State Nigeria. The data are on reported cases of children treated, survived and did not survive of pneumonia for 2007-2016 as presented in Appendix Table 1. To facilitate easy analysis, some notations compactable to r-statistical software were implemented as presented in Appendix 2. Adopting notation of time normalization in Appendix 2, data presented are analyzed based on the estimates of proportion surviving at time (t) using the Kaplan-Meier method. The results are shown in Table 1.

RESULTS AND DISCUSSION

From Table 1, covariate of sex 1 represents males while sex 2 represents females. Time factor in survival analysis has varying values. Since time is

Table 1. Time and incidence of pneumonia at ISTH.

Year	Survival	Death	ratio (ϖ)	Status	Sex	Time(days)
2007	292	18	16.222	1	1	112.500
2008	266	15	17.733	1	1	102.914
2009	260	24	10.833	1	1	168.462
2010	255	25	10.200	1	1	178.922
2011	215	33	6.515	0	1	280.116
2012	206	45	4.578	0	1	398.665
2013	239	44	5.432	0	1	335.983
2014	205	31	6.613	0	1	275.976
2015	235	43	5.4651	1	1	282.284
2016	208	26	8.000	0	1	228.125
2007	361	14	25.786	1	2	70.776
2008	281	18	15.611	1	2	116.904
2009	257	25	10.280	1	2	177.529
2010	182	27	6.741	0	2	270.742
2011	250	27	9.259	1	2	197.100
2012	190	24	7.917	0	2	230.526
2013	309	24	12.875	1	2	141.748
2014	234	27	8.667	1	2	210.577
2015	246	24	10.250	1	2	162.222
2016	249	31	8.032	0	2	227.209

all through the years, transformation of variables and effects was carried out for optimum performance. Time and incidence table was obtained by transformation technique which was generated using r-statistical fit *xlab* for patient time (years/days) and *ylab* for survival probabilities. The Kaplan-Meier curve was used to estimate percentiles survival distribution of median time and mean time in years; the survival distribution showing the 95% confidence interval, standard error, survival rate, and number at risk at a particular event is displayed in Table 2 and the survival probability is presented in Table 3.

Table 2. Survival distribution of pneumonia cases.

Time	Risk	Event	Survival	Std. Error	Lower C.I.	Upper C.I.
70.8	20	1	0.95	0.0487	0.859	1
102.9	19	1	0.9	0.0671	0.778	1
112.5	18	1	0.85	0.0798	0.707	1
116.9	17	1	0.8	0.0894	0.643	0.996
141.7	16	1	0.75	0.0968	0.582	0.966
162.2	15	1	0.7	0.1025	0.525	0.933
168.5	14	1	0.65	0.1067	0.471	0.897
177.5	13	1	0.6	0.1095	0.42	0.858
178.9	12	1	0.55	0.1112	0.37	0.818
197.1	11	1	0.5	0.1118	0.323	0.775

From Table 2 it is observed that at days (time) 70.8 the chances of survival were 0.95; but as the days (time) increases to 197.1, the chances of survival dropped to 0.5. This indicates that as time progresses the chances of survival decrease drastically as can also be seen in Figures 1 and 2. As the length of time increases, survival probability drops as shown in Figures 1 and 2. The Kaplan-Meier estimate is in solid line and its 95% confidence intervals are in dotted lines as indicated in Figure 1. From Figure 2, line 1 (blue color) represents the male sex covariates

Table 3. Survival probability of pneumonia cases.

Year	Sex	Reported cases	Survivals	Survival probability
2007	1	310	292	0.941935
2008	1	281	266	0.946619
2009	1	284	260	0.915492
2010	1	280	255	0.910714
2011	1	248	215	0.866935
2012	1	251	206	0.820717
2013	1	283	239	0.844522
2014	1	236	205	0.868644
2015	1	278	235	0.845323
2016	1	234	208	0.888888
2007	2	375	361	0.962666
2008	2	299	281	0.939799
2009	2	282	257	0.911347
2010	2	209	182	0.870813
2011	2	277	250	0.902527
2012	2	214	190	0.887850
2013	2	333	309	0.927927
2014	2	261	234	0.896551
2015	2	270	246	0.911111
2016	2	280	249	0.889285

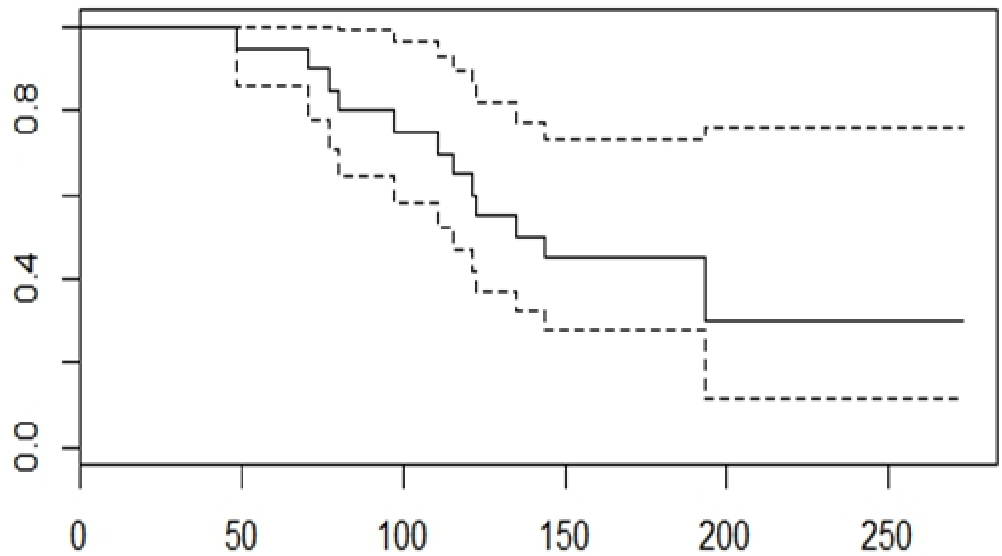

Figure 1. The Kaplan -Meier curve of pneumonia patients.

in Group 1 while line 2 (red color) represents the female sex covariates in Group 2. The trend shows that the proportion of under-five male children surviving under pneumonia is higher than that of the females in the transformed time. From the curves, the horizontal lines represent the survival duration for the interval while the height of the vertical lines shows the change in cumulative probability; censored observation are indicated by tick marks. This helps to reduce the cumulative survival between the intervals. For each time interval, survival probability is the proportion of patients that survive beyond a specified time. These estimates of survival probability are frequently referred to as reliability estimates. It is calculated as the number of patients surviving divided by the number of patients at risk. The survival time follows an exponential distribution with mean time of 160T (2.19 years), median value of

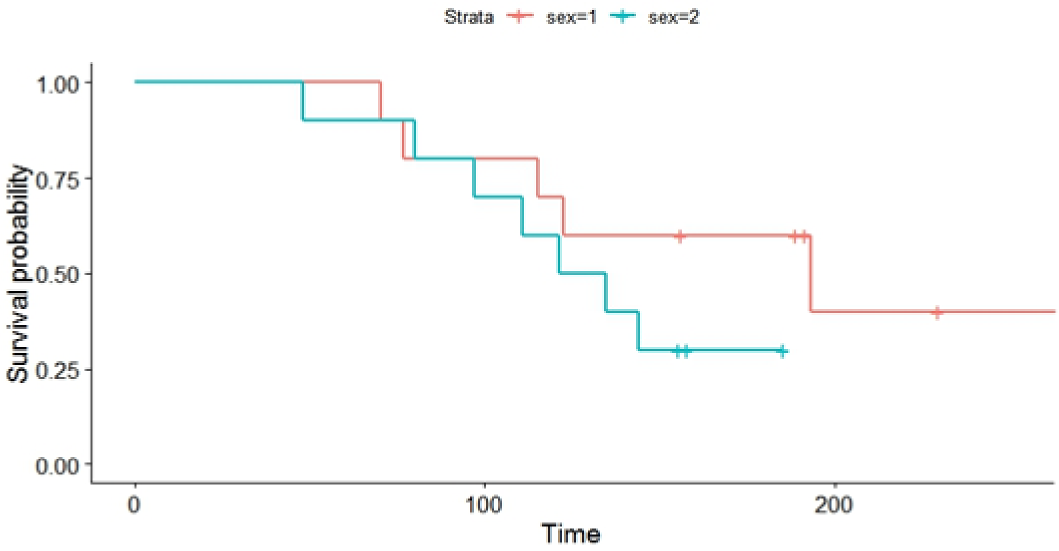

Figure 2. Survival time of pneumonia patients.

165.35T (2.26years) and 95% confidence interval (C.I 2.147-2.373) (Table 4). The log rank test shows that there is statistical significant evidence of the survival distributions between the male and female respondents (p-value =0 .03 < 0.05) (Table 5).

Table 4. Survival time of patients.

Mean time	C.I at 95%	Median time 50%	P<value (Sig. Diff.)
160T (2.19 years)	2.147-2.373	165.35T (2.26 years)	P= 0.03 < 0.05

Table 5. Log rank test of patients.

	N	Observed	Expected	(O-E)^2/E	(O-E)^2/V
Sex = Male	10	5	6.87	0.509	1.28
Sex = Female	10	7	5.13	0.681	1.28

Chisquare = n-1 degree of freedom, p = 0.03.

With a p-value 0.03 < 0.05, there is a statistical significant difference in the survival distributions between the male and female respondents; this negates the null hypothesis. This indicates that the expected mean surviving rate of male children age under-five (6.87) is significantly higher than that of the female children age under-five (5.13) with a p-value of 0.03. However, Table 6 indicates that the mean survival rates of both male and female are 11.32216 and 8.60714 respectively, and the mean survival rate (bar) as presented in Table 7 indicates that male is 1.132216, while female is 0.860714 within the given time period. This goes to indicate that the mean survival rate (bar) proportion of male patients surviving is higher than that of the females.

Results obtained from the analysis showed that at 70.8 days (time), the chance of survival was 0.95 but with the length of time at 197.1 the chances of survival decreased to 0.5. This trend reveals that as the length of time increases the proportion of surviving decreases, indicating an increase in mortality of infants under five (0-5 years) over time. The Kaplan-Meier estimate is in solid line and its 95% confidence intervals are in dotted lines as indicated in Figure 1. From Figure 2, line 1 (blue color) represents the male sex covariates in group 1 while line 2 (red color) represents the female sex covariates in group 2.

Table 6. Mean survival rate for male and female patients.

Year	Male	Treated cases	Deaths cases	MSR (%)	Female	Treated cases	Deaths cases	MSR (%)
2007	1	310	18	5.80645	2	375	14	3.73333
2008	1	281	15	5.33807	2	299	18	6.02006
2009	1	284	24	8.4507	2	282	25	8.86524
2010	1	280	25	8.92857	2	209	27	12.9187
2011	1	248	33	13.3065	2	277	27	9.74729
2012	1	251	45	17.9283	2	214	24	11.215
2013	1	283	44	15.5477	2	333	24	7.2072
2014	1	236	31	13.1356	2	261	27	10.3448
2015	1	278	43	15.4676	2	270	24	8.88888
2016	1	234	26	11.1111	2	280	31	11.0714
		2,685	304	11.3222		2,800	241	8.60714

Table 7. Mean Survival Rate (Bar).

Sex	MSR	N (2007-2016)	MSR (Bar)
Male	11.32216	10	1.132216
Female	8.60714	10	0.860714

The trend shows that the proportion of under-five male children surviving pneumonia is higher than that of the females in the transformed time. The survival time follows an exponential distribution with mean time of 160T (2.19 years), median value of 165.35T (2.26years) and 95% confidence interval (C.I 2.147-2.373) as shown in Table 4. The log rank test shows that there is statistical significant evidence of the survival distributions between the male and female respondents (p-value =0.03 < 0.05) (Table 5). However, Table 6 indicates that the mean survival rate of both male and female children are 11.32216 and 8.60714 respectively, and the mean survival rate (bar) as presented in Table 7 indicates that male is 1.132216, while female is 0.860714 within the given time period. This goes to indicate that the mean survival rate (bar) proportion of male patients is higher than that of the females.

Conclusion

The Kaplan-Meier survivorship estimates was used to examine the model fit. The curve checks whether the observed number of events is significantly different from the expected number of events in groups differentiated by risk scores. Results obtained from the analysis showed that at 70.8 days (time), the chance of survival was 0.95 but with the length of time at 197.1 the chances of survival decreased to 0.5. This trend reveals that as the length of time increases the proportion of surviving decreases indicating an increase in mortality of infants under five (0-5 years) over time as a result of the effect of pneumonia. The log rank test shown in Table 4 indicates significant difference (p = 0.03<0.05) between male and female respondents. The survival time for the hospital follows an exponential distribution with mean time of 2.19years, while the median survival time was estimated to be 165.35T (2.26years, with its 95% confidence interval of 2.147-2.373 years).

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Adetoro, G.W. and Amoo, E.O. (2014). A statistical analysis of child mortality Evidence from Nigeria Journal of Demography and Social Statistics; 1: 110 - 120.

Ahmed, F.E, Vos, D.W and Holbert, D. (2007). Modeling survival in colon Cancer a

Methodological Review. Statistical primer basics of survival analysis; 6:15-21.

Baulies, S, Belin, L and Mallon P. (2015). Time - Varying effect and Long term survival analysis in breast cancer patients. *Journal of Applied Sciences;* 6: 113-130.

Dafni, U. (2011). Landmark analysis. *Journal on cardiovascular outcomes*; 4: 363-371.

Finlay, E, Ozaltin, J and Canning F. (2011). Association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemic for first births, evidence from 55 low and middle income countries. *BM journal;* 1: 2-8.

George, B, Seals, S and Aban, I. (2014). Survival analysis and regression models *Journal on Cardiological Studies;* 21: 686 – 694.

Hosmer, D.W, Lemeshow, S and May S. (2008). Applied Survival. Analysis Regression modelling of time to event data. Wiley; 34: 1352 - 1361.

Irrua Specialist Teaching Hospital (2017). ISTH Child Health Records Department Annual Report, Edo State. Nigeria.

Katrona, P and Katona A.J. (2008). The Interaction between nutrition and infection. *Clinical infectious Diseases Oxford Journal;* 46: 1582-1588.

Kleinbaum, G and Klein, M. (2012). Survival Analysis. 3rd Edition, New York Springer. A self-learning text 3:31-37.

Kuti, B. P and Oyelami, A.O. (2015). Childhood community acquired pneumonia at the Wesley guild hospital Ilesa, Prevalence, pattern and outcome determinants. *Niger Journal on Health Sciences;* 15: 98-104.

Ogbonaya, R and Aminu, M. (2009). Nigeria North-West Battling malnutrition, Child and maternal mortality. This Day, all Africa.com: retrieved 07/06/2018.

Ogunjimi, L.O, Ibe, R &Ikorok, M.M. (2012). Curbing maternal and child mortality: The Nigerian experience.*Experience International Journal of Nursing and Midwifery;* 4:33-39.

Okumale, O. (2017). Pneumonia as a major killer disease of infants in Nigeria. *Journal on Perfect Health Initiative; 2: 21-28.*

Simona, D. (2008). An introduction to survival analysis – my courses. *Cornell statistical consulting unit stat news;* 78: 1- 3

Udofia, I. and Okonofua, F. (2008). Preventing primary post-partum hemorrhage in unskilled births in Africa. *Journal of Africa Reproductive Health;* 12: 7- 9.

Appendix Table 1. Data of reported cases of children treated, survived and died of pneumonia in Irrua Specialist Teaching Hospital (ISTH), Irrua.

Infants age group	Reported cases males	Survival males	Death recorded males	Reported cases females	Survivals females	Death recorded female
0-28DAYS	121	116	5	151	148	3
29-364 DAYS	86	81	6	92	87	5
1-4 YEARS	66	62	4	71	68	3
4+ 5 YEARS	37	33	3	61	58	3
TOTAL 2007	**310**	**292**	**18**	**375**	**361**	**14**
0-28DAYS	101	98	3	132	125	7
29-364 DAYS	93	88	5	83	79	4
1-4 YEARS	39	37	2	47	46	1
4+ 5 YEARS	48	43	5	37	31	6
TOTAL 2008	**281**	**266**	**15**	**299**	**281**	**18**
0-28DAYS	127	121	6	115	111	4
29-364 DAYS	87	78	9	94	87	7
1-4 YEARS	31	28	3	42	36	6
4+ 5 YEARS	39	33	6	31	23	8
TOTAL 2009	**284**	**260**	**24**	**282**	**257**	**25**
0-28DAYS	111	107	4	83	76	7
29-364 DAYS	75	68	7	58	53	5
1-4 YEARS	51	43	8	26	20	6
4+ 5 YEARS	43	37	6	42	33	9
TOTAL 2010	**280**	**255**	**25**	**209**	**182**	**27**
0-28DAYS	102	93	9	103	94	9
29-364 DAYS	82	71	11	92	85	7
1-4 YEARS	35	27	8	36	30	6
4+ 5 YEARS	29	24	5	46	41	5
TOTAL 2011	**248**	**215**	**33**	**277**	**250**	**27**
0-28DAYS	99	88	11	93	87	6
29-364 DAYS	68	53	15	33	26	7
1-4 YEARS	48	34	14	47	39	8
4+ 5 YEARS	36	31	5	41	38	3
TOTAL 2012	**251**	**206**	**45**	**214**	**190**	**24**
0-28DAYS	104	90	14	131	125	6
29-364 DAYS	82	70	12	79	71	8
1-4 YEARS	35	23	12	45	39	6
4+ 5 YEARS	62	56	6	78	74	4
TOTAL 2013	**283**	**239**	**44**	**333**	**309**	**24**
0-28DAYS	111	102	9	104	97	7
29-364 DAYS	58	49	9	92	84	8
1-4 YEARS	36	29	7	25	17	8
4+ 5 YEARS	31	25	6	40	36	4
TOTAL 2014	**236**	**205**	**31**	**261**	**234**	**27**
0-28DAYS	118	103	15	110	104	6
29-364 DAYS	69	58	11	82	73	9
1-4 YEARS	29	18	11	31	25	6
4+ 5 YEARS	62	56	6	47	44	3
TOTAL 2015	**278**	**235**	**43**	**270**	**246**	**24**

Appendix Table 1. Continue

0-28DAYS	101	93	8	131	123	8
29-364 DAYS	73	67	6	66	55	11
1-4 YEARS	28	21	7	43	36	7
4+ 5 YEARS	32	27	5	40	35	5
TOTAL 2016	**234**	**208**	**26**	**280**	**249**	**31**

Source: Records Department ISTH (2017).

Appendix 2. Time normalization notations.

$$Surv-death = \varpi = \frac{survival_i}{death_j} \text{.................}(a), \textit{where i=survival at } i-th \textit{ term, and } j\textit{=death at } j-th \textit{ term}$$

$$death-surv = \varpi' = \frac{death_j}{survival_i} \text{................}(b), \textit{where i=survival at } i-th \textit{ term, and } j\textit{=death at } j-th \textit{ term}$$

$$Summation\ Ratio = \psi = \frac{\sum_{i=1}^{n} survival}{\sum_{j=1}^{m} death},\ i = 1...n;\ j = 1...m \text{..}(c)$$

$$Summation\ Penalty(\lambda) = \psi - 10\%\psi \Rightarrow \lambda = \psi - 0.1*\psi \text{...}(d)$$

$$T = 365 * 5 = 1825,\ hence,$$

$$time(t_j) = \varpi' * T \text{..}(e)$$

$$Status = \begin{cases} 1, & if\ \varpi > \lambda \\ 0 & otherwise \end{cases} \text{..}(f)$$

$$rescaled\ time(t'_j) = t_j \times \frac{5}{365} \text{..} (g)$$