International Journal of Innovative Mathematics, Statistics & Energy Policies
12(4):6-14, Oct-Dec., 2024

 www.seahipublications.org **ISSN: 2467-852X**

DOI: 10.5281/zenodo.13983872

Statistical Analysis Of Infant Mortality Using Cox Proportional Hazard Regression Case Study Of Taraba State

Kamalu Ibrahim Balansana, Kazeem E. Lasisi, Mohammed Salawu Atureta, Musa Usman Bawa, Danjuma Abdulazeez & Hafiza Inusa Idris.

Department of Statistics, Modibbo Adama University, Yola, Adamawa State, Nigeria
Department of Statistics, Abubakar Tafawa Balewa University, Bauchi, Nigeria
Department of Mathematics, Abubakar Tafawa Balewa University, Bauchi, Nigeria
Department of Statistics, Federal Polytechnic Damaturu, Yobe State, Nigeria
Department of Statistics, Federal Polytechnic Shendam, Plateau State, Nigeria

Corresponding Author: Kamalu Ibrahim Balansana,
Email: kamaluibrahim5@gmail.com

ABSTRACT
Infant mortality remains a significant public health challenge globally, with Nigeria experiencing one of the highest rates in the world. This study analyses the determinants of infant mortality in Taraba State, Nigeria, using Cox Proportional Hazard Regression. The data, obtained from the Federal University Wukari Teaching Hospital, spans the years 2012 to 2022 and includes key variables such as maternal age, marital status, educational level, birth weight, delivery method, and place of birth. The study aims to identify key factors influencing infant mortality and estimate the parameters associated with infant mortality in model Cox Proportional Hazard Regression focuses on the survival time of infants. Results indicate that marital status, educational level, and birth weight are significant predictors of infant mortality in Taraba State. Gender, maternal age, place of birth, and method of delivery do not show significant influence on the hazard rate or mortality likelihood. The study concludes that marital status, educational level and birth weight are the influencing factors of infant mortality and recommends incorporating environmental and socio-economic factors in future studies to provide a more comprehensive understanding of infant mortality determinants in the region.
Keywords: Infant Mortality, Cox Proportional Hazard Model, Birth Weight, Survival Analysis

INTRODUCTION
Infant mortality (IM) is the death of an infant before his or her first birthday. (CDC 2022) Infant mortality rate (IMR) is a key indicator of the overall health of a society and is essential for social and economic development. (CDC 2022) In the National Demographic Health Survey (NDHS), IMR of each year was calculated as the number of infant deaths among the total number of alive-born children in the year.
This mortality rate is used more widely as a rough gauge of a community's socioeconomic condition, poverty levels, and general health status in addition to reflecting the risk of infant death. (Child Health Program. 2019). Studies show that there are several determinants of IM in Nigeria. In 2010, low birth weight was highlighted as the most common cause of IM accounting for 25% of IM. The study also identified lack of delivery attendants, home delivery and traditional birth attendants as predictors of IM in Nigeria.

Nigeria reported 72 deaths per 1000 live births among infants in 2020 with disparities across its regions and geopolitical zones. (WB 2022) Nigeria has worse IMR compared with neigh boring West African countries such as Benin, Cameroon, Togo and Ghana, with 57, 48, 44 and 33 deaths per 1000 live births, respectively. (UNICEF 2022) The rate makes Nigeria one of the countries with the highest IMR rate.

In 2019, globally, the estimated IMR was 35 per 1000 live births (LB), whereas in Sub-Saharan Africa, it was 64 per 1000 LB. (UNICEF 2019). Similarly, Sub-Saharan Africa has seen the least decline in IMR and under-five deaths (UN 2018). Ethiopia is one of the few nations with the highest infant mortality rates. (Weldearegawi, et al. 2018).

Pakistan is the 5th most populated developing country in the world and placed 2nd in Islamic countries after Indonesia, with a higher proportion lived in rural areas, lower literacy rate, insufficient health care settings, and poor quality of life. Pakistan is among the countries that contribute to 50% of all new-born deaths (UN IGME 2017). In South Asia, Pakistan placed at the top with the highest IMR (64/1,000 live births) followed by the neighbouring border sharing country Afghanistan (53/1,000), India (35/1,000), Bangladesh (28/1,000), Nepal (28/1,000), Iran (13/1,000) and Sri Lanka (8/1,000) (Gallup 2018).

Despite the implementation of a set of social and health policies that positively affected the health of population (Gomes, T.G.A.C.B.et al.2016), since 2009, Brazil has been experiencing a slower decline in infant mortality (UNICEF 2017) that has remained at high levels and presents significant regional disparities. In 2016, the country recorded an increase in the mortality of children under one and under five years old, which disrupted a 25-year downward trend (UNICEF 2017).

In West African countries, there was a high incidence of killer diseases which affected maternal and infant health leading to high mortality of infants. This can be attributed to the type of vegetation in West Africa and other factors which include: the sanitary condition of our local populace, mode of feeding, and also lack of health infrastructures. All these created problems that led to high infant mortality (CDC, 2020).

Again, other causes of infant mortality are recognized killer childhood diseases, such as measles, malaria, smallpox, tetanus, whooping cough, hepatitis, poliomyelitis, diphtheria, and jaundice. In the late 80s, the government of Nigeria, through the National Program on Immunization and Expanded Program on Immunization, achieved a lot regarding the improvement of infant health (CDC, 2020). This is supported by Samantha Slinkard et al., 2018 who reported lack of access to Antenatal care (ANC) or delayed ANC initiation as an important risk factor for increased IM.

According to the 2019 NDHS data, just 18% of women who received ANC did so during the first trimester, and 34% of women did not receive ANC at all. Nigeria has a high IMR due to a number of factors, including the mother's age, her socioeconomic position, and her region, all of which affect the impact of ANC start. The location of the child's birth, the skill of the birth attendant, caesarean delivery, the sequence and spacing of the births, the mother's age and education, and the wealth index are other factors that have been highlighted.

Even though significant improvements have been made in the quality and access to neonatal and infant care during the past decade, large educational, socioeconomic, racial, ethnic, geographic and behavioural disparities persist, and appear to be responsible for significant disparities in IMR among different subgroups. Certain maternal and infant characteristics have important associations with IMR, and this study attempted to quantify major maternal and infant predictors and trace associated mortality trends during the study period. The neonatal mortality rate (NMR) is a standardized measurement of deaths in the first 27 days of life per thousand live births and has different significance. NMR reflects prevalence of antenatal abnormalities such as birth defects and intrauterine infections, as well as quality of neonatal care. This rate is less sensitive to educational and socioeconomic differences in the population (Samantha Slinkard et al., 2018)

Infant deaths are intrinsically linked to several causes and factors which need extensive studies. Without addressing this, it's going to be hard to see the same level of progress that has been made in the reduction of infant mortality in last decade. Cause leads directly to a death, while a contributor is a risk factor that makes the death more likely to occur (NICHHD 2016).

Cause of infant death information are critical for formulating good public health policy (James SL. et al 2019). Emphatically, detecting and generating empirical information on cause of infant death and patterns of association factors on cause-specific mortality at population level is the first priority and basically utmost essential to take evidence-based measures. This ultimately contributes in reducing infant morbidity and mortality. However, determination of causes of death is a global concern as around half of the world's children

deaths pass without any formal registration of the cause of death (CR 2020). Such situation leads to a major problem for understanding the infant's health problems in general and difficult for planning the best solutions.
In recent times, it has been a major worldwide health issue facing almost every nation. Poverty and ignorance are the basic causes of the high death rate among newly born and small-aged children worldwide, especially in third-world nations. The basic indicators of social development, human being and health status are infant mortality (Ahmad et al., 2000).
It is time to take serious steps to overcome such an important issue for both developed and developing nations of the world and make effective health policies. The infant mortality rate is high among the poor nation of the world compared to those nations with high per capita income. Lack of resources and unequal distribution of wealth are other causes of the high death rate among the less developed nation. Four to five million children die every year within the first month of their birth. This shows that every minute, seven to eight newly born children lose their life (WHO, 2016).
Despite the use of sophisticated statistical technique (logistic regression) by Damilola et al (2022) in modelling the infant mortality rate. The incident in Taraba state keeps on increasing as infant are dying in large number, this calls for a need to look for a competitive model that will handle the problem better hence Cox Proportional Hazard model will be consider in the study. The thesis focuses on analysing infant mortality in Taraba State, Nigeria, using a statistical model Cox Proportional Hazard Model. The study aims to identify key factors influencing infant mortality and estimate the parameters associated with infant mortality in model.
Despite being totally preventable and treatable, common infectious diseases are still killing younger children in large numbers, according to the Centre for Disease Control and Prevention (CDC, 2020).

MATERIAL AND METHOD

The research adopted a Secondary data from the Federal University Wukari Teaching Hospital, covering 2012-2022, was used. Key variables include maternal age, marital status, education, birth order, birth weight, delivery method, and place of birth. The study applies Cox Regression to estimate the effects of the variables on infant mortality.

The Cox's Proportional Hazard

The Cox Proportional Hazard (assuming hazard is constant for all patients) model is a semi parametric model in which the hazard function of the survival time is given by

$$h(t,\boldsymbol{X}) = h_0(t)exp\left[\sum_{i=1}^{p} \beta_i X_i\right] = h_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{X}) \ldots\ldots\ldots\ldots\ldots.\ldots\ldots\ldots\ldots.\ldots 1$$

$$h(t,\boldsymbol{X}) = h_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.\ldots.2$$

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., $\boldsymbol{X} = (x_1, x_2, \ldots, x_p)'$ is the values of vector of explanatory variables for a particular individual, and $\boldsymbol{\beta}' = (\beta_1, \beta_2, \ldots, \beta_p)$ is a vector of regression coefficients.
The corresponding survival functions are related as follows:

$$S(t,\boldsymbol{X}) = [s_0(t)]exp\left[\sum_{i=1}^{p} \beta_i X_i\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..3$$

This model, also known as the Cox regression model, 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 a semi-parametric model. The beauty of the Cox approach is that vagueness creates no problems for estimation.
Even though the baseline hazard is not specified, we can still get a good estimate for regression coefficients β hazard ratio, and adjusted hazard curves. The measure of effect is called hazard ratio. The hazard ratio of two individuals with different covariates x and x^* will be given by:

$$HR = \frac{\hat{h}(t,\boldsymbol{X}^*)}{\hat{h}(t,\boldsymbol{X})} = \frac{\hat{h}_0(t)\exp(\sum_{i=1}^{p}\hat{\beta}_i X_i^*)}{\hat{h}_0(t)\exp(\sum_{i=1}^{p}\hat{\beta}_i X_i)} = \exp\left[\sum_{i=1}^{p}\hat{\beta}_i\,(X_i^* - X_i)\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4$$

For the sake of interpretation, we usually want $HR \geq 1$ i.e., $\hat{h}(t,\boldsymbol{X}^*) \geq \hat{h}(t,\boldsymbol{X})$
This hazard ratio is time-independent, which is why this is called the proportional hazards model.

Model Specification
Taking survival time as the dependent variable and mother age, marital status, level of education, order of birth, babies' weight, method of birth, place of birth. as independent prognostic factors, and based on the fact that the ratio of the hazards for any two individuals *i* and *j* is constant over time the expected model will assume the form:

$$S(t,\boldsymbol{X}) = [s_0(t)]^{exp(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5$$

Where β_i's are estimated coefficients of the regression model such that, $\beta_0 = \ln(s_0)$ *, is the baseline constant for the regression model*

β=relative risks, X_1 = mothers' age, X_2 = marital status, X_3 = level of education, X_4= order of birth, X_5 = babies' weight, X_6 = method of birth, X_7 = place of birth.

These variables in SPSS are represented by using dummy variables (1 or 0) except the order of birth, babies' weight and survival time. The required model for the research is however, dependent on the significance of each of these factors at a level of significance of 0.05.

RESULT AND DISCUSSION

Analysis
Data were analyzed using the statistical package for social sciences (SPSS) software version 22.

4.2.1 Cox Regression

Table 4.1 Case Processing Summary

		N	Percent
Cases available in analysis	Event[a]	403	24.9%
	Censored	1217	75.1%
	Total	1620	100.0%
Cases dropped	Cases with missing values	0	0.0%
	Cases with negative time	0	0.0%
	Censored cases before the earliest event in a stratum	0	0.0%
	Total	0	0.0%
Total		1620	100.0%

a. Dependent Variable: Duration

Table 4.2 Categorical Variable Codingsa,c,d,e,f,g

		Frequency	(1)	(2)
GENDER[b]	1=MALE	829	1	
	2=FEMALE	791	0	
AGE[b]	1=<25	391	1	0
	2=25-40	1194	0	1
	3=40 ABOVE	35	0	0
Mstatus[b]	1=MARRIED	1516	1	
	2=SINGLE	104	0	
Elevel[b]	1=PRIMARY	363	1	0
	2=SECONDRY	847	0	1
	3=TERTIARY	410	0	0
Mdelivery[b]	1=SVD	1306	1	
	2=C/S	314	0	
Pbirth[b]	1=HOSPITAL	1616	1	
	2=HOME	4	0	

a. Category variable: GENDER
b. Indicator Parameter Coding
c. Category variable: AGE
d. Category variable: Mstatus
e. Category variable: Elevel
f. Category variable: Mdelivery
g. Category variable: Pbirth

Block 0: Beginning Block

Omnibus Tests of Model Coefficients

-2 Log Likelihood
5853.208

Block 1: Method = Enter

Table 4.3 Omnibus Tests of Model Coefficients[a]

	Overall (score)			Change From Previous Step			Change From Previous Block		
-2 Log Likelihood	Chi-square	Df	Sig.	Chi-square	Df	Sig.	Chi-square	df	Sig.
5587.130	298.764	10	.000	266.078	10	.000	266.078	10	.000

a. Beginning Block Number 1. Method = Enter

Table 4.4 Variables in the Equation

							95.0% CI for Exp(B)	
	B	SE	Wald	Df	Sig.	Exp(B)	Lower	Upper
GENDER	.145	.101	2.052	1	.152	1.156	.948	1.409
AGE			2.514	2	.284			
AGE(1)	.754	.477	2.502	1	.114	2.126	.835	5.411
AGE(2)	.704	.457	2.376	1	.123	2.022	.826	4.947
Mstatus	-.770	.167	21.231	1	.000	.463	.334	.643
Elevel			38.489	2	.000			
Elevel(1)	-.950	.219	18.825	1	.000	.387	.252	.594
Elevel(2)	.212	.130	2.671	1	.102	1.236	.959	1.594
ORbirth	.039	.039	1.035	1	.309	1.040	.964	1.121
Bweigth	-.976	.075	168.139	1	.000	.377	.325	.437
Mdelivery	-.031	.122	.064	1	.800	.970	.763	1.232
Pbirth	-.876	.517	2.877	1	.090	.416	.151	1.146

Table 4.5 Covariate Means

	Mean
GENDER	.512
AGE(1)	.241
AGE(2)	.737
Mstatus	.936
Elevel(1)	.224
Elevel(2)	.523
ORbirth	2.852
Bweigth	2.944
Mdelivery	.806
Pbirth	.998

RESULT

From table 4.5 show that **Gender (male):** coefficient (B) 0.145 with p-value 0.152 and hazard ratio (Exp(B)) 1.156**,** male have a 15.6% higher hazard than females, but this is not statistically significant. **Age(1) < 25:** coefficient (B) 0.754 with p-value 0.114 and hazard ratio (Exp(B)) 2.126, **Age(2)25-40:** coefficient (B) 0.704 with p-value 0.123 and hazard ratio (Exp(B)) 2.022, about the two ages have twice the hazard and are not significance. **Marital Status:** coefficient (B) -0.770 with p-value 0.001 and hazard ratio (Exp(B)) 0.463, married individuals have a 53.7% lower hazard than single and is significant. **Educational Level: Primary** the coefficient (B) -0.950 with p-value 0.001 and hazard ratio (Exp(B)) 0.387, primary educated have a 61.3% lower hazard than those with secondary or tertiary educations and is significant**. Secondary** the coefficient (B) 0.212 with p-value 0.102 and hazard ratio (Exp(B)) 1.236, secondary educated have a 23.6% higher hazard than tertiary and is not significant. **Place Of Birth:**

coefficient (B) -0.876 with p-value 0.090 and hazard ratio (Exp(B)) 0.377, those born in hospitals have a 58.4% lower hazard but not statistically significant. **Birth Weight** coefficient (B) -0.976 with p-value 0.001 and hazard ratio (Exp(B)) 0.377, lower birth weight is associated with a 62.3% lower hazard. **Method Of Delivery:** coefficient (B) -0.031 with p-value 0.800 and hazard ratio (Exp(B)) 0.970, no significant difference in hazard based on method of delivery.

SUMMARY:

Marital Status being married significantly reduces the hazard rate (risk of event) compared to being single. **Primary Education** individuals with only primary education have a significantly lower hazard compared to those with secondary or tertiary education. **Birth Weight** lower birth weight is significantly associated with a lower hazard. While **Gender, Age, Place Of Birth And Method Of Delivery** do not significantly affect the hazard rate in this analysis. From the result marital status, educational level and birth weight are the influencing factors of infant mortality.

FINDINGS:

Marital status, educational level, and birth weight were significant factors in the model.
Gender, maternal age, place of birth, and method of delivery were not statistically significant.

CONCLUSIONS

Marital status, educational level, and birth weight are key predictors of infant mortality in Taraba State. The study recommends considering environmental factors in future research for a more comprehensive understanding of infant mortality.

REFERENCES

Agbedeyi Odimientimi Desmond, Edike Nnamdi Adigwe Rosemary Umeadi and Tuggen O. Edward,(2023) STATISTICAL ANALYSIS OF INFANT MORTALITY RATE IN NIGERIA *global scientific journal GSJ: Volume 9, Issue 8, August 2021, online: ISSN 2320-9186*

Ahmad, O.B., Lopez, A.D., Inoue, M., (2000). The decline in child mortality *a reappraisal. Bull. World Health Organ. 78, 1175–1191*

Aminu, I. (2017); "Infant Mortality Analysis of a Tertiary Hospital in Adamawa State, Nigeria".(Unpublished undergraduate project from the department of Statistics and Operations Research, MAUTECH).

Association of Maternal and Child Health Program (2019). Infant Mortality Toolkit: A standard Approach for Examining Infant Mortality. *Association of Maternal and Child Health Program*.

Berhe Weldearegawi, Yohannes Adama Melaku, Semaw Ferede Abera, et al (2015) Infant Mortality and causes of infant deaths in rural Ethiopia: a population-based cohort of 3684 births *BMC Public Health (2015) 15:770 DOI 10.1186/s12889-015-2090-x*

Centre for Disease Control and Prevention (2020); Infant mortality
https://www.cdc.gov/reproductivehealth/maternalinfanthealth/infantmortality.htm

Chowdhury Q. H, Rafiqul I. R, Hossain K. 2010. Socioeconomic determinants of neonatal, post neonatal, infant and child mortality. *International Journal of Sociology and Anthropol, 2:118-125*

Civil registration (2020) why counting births and deaths is important [Internet].
https://www. who.int/news-room/fact-sheets/detail/civil-registration-why-counting-births-and-deaths-is-important.

Damilola M. Shobiye, et al (2022) Infant mortality and risk factors in Nigeria in 2013−2017: *A population-level study eClinical Medicine 2022;51: 101622*

Eunice Kennedy Shriver National Institute of Child Health and Human Development (2016). What causes infant mortality? [Internet]. NICHHD Reviewed; 2016 [cited 2020 May 6] *https://www.nichd.nih.gov/ health/topics/infant-mortality/topicinfo/causes*

Ezeh OK, Agho KE, Dibley MJ, et al. (2015) Risk factors for postneonatal, infant, child and under-5 mortality in Nigeria: a pooled cross-sectional analysis. *BMJ Open 2015;5: e006779. doi:10.1136/bmjopen-2014- 006779*

Gallup. (2018) Gallup report of infant mortality Pakistan; Available *from:http://gallup.com.pk/wp-content/ uploads/2018/04/Infant-mortality-PR-1.pdf.*

Gomes, T.G.A.C.B., Ferreira, L.D.S., Queiroz, M.N.D., Reis Netto, Costa, A.M et al. (2016) *Mortalidade na infância no Brasil e regiões no período de 2000 a 2011: O impacto da atenção básica. Comun. Ciênc. Saúde 2016, 27, 259–266. [CrossRef]*

James SL, Lozano R, Freeman MK, et al. (2019) Performance of InterVA for assigning causes of death to verbal autopsies: Multisite validation study using clinical diagnostic gold standards. *Popul Health Metrics. 2016; 9(50). https://doi.org/10.1186/1478-7954-9-50 PMID: 21819580*

Kumar P. P, File G. 2010. Infant and Child Mortality in Ethiopia: *A Statistical Analysis Approach. Ethiopian Journal of Education and Science 5 (2): 51-7.*

Lozano R, Freeman MK, James SL. et al. (2016) Performance of InterVA for assigning causes of death to verbal autopsies: Multisite validation study using clinical diagnostic gold standards. *Popul Health Metrics. 2016; 9(50). https://doi.org/10.1186/1478-7954-9-50 PMID: 21819580*

Mojekwu, J. N. and Ajijola, L. A. 2011. Developing a Model for Estimating Infant Mortality rate of Nigeria. *Journal of Research in International Business and Management 1(2): 164-170.*

Okorie C.E., Balansana K.I., Bulus S.M. (2023), Statistical Analysis of Infant Mortality Rate: A Case Study of Taraba State. *African Journal of Mathematics and Statistics Studies 6(3), 48-58. DOI: 10.52589/AJMSS-YIBS555O*

Paul Andrew Bourne, Subrina South, Sannekia Ellis, Donae Hansle, et al (2022) An Assessment of Infant Mortality Rates in Jamaica, 2002-2021. International Journal of Interdisciplinary Invention & Innovation in Research *https://eurekajournals.com/IIIRJ.html ISSN: 2581-6438*

Rai SK, Kant S, Srivastava R, et al. (2017) Causes of and contributors to infant mortality in a rural community of North India: evidence from verbal and social autopsy. *BMJ Open 2017;7: e012856. doi:10.1136/ bmjopen-2016-012856*

Samuel Mebrahtom, Worku A, Gage DJ (2022) Causes of infant deaths and patterns of Associated factors in Eastern Ethiopia: Results of verbal autopsy (InterVA-4) study. PLoS ONE 17(8): e0270245. *https://doi.org/10.1371/journal. pone.0270245*

Samantha Slinkard, JR Pharr, T Bruno, et al (2018) Determinants of infant mortality in Southeast Nigeria: results from the healthy beginning initiative, 2013-2014*International Journal of MCH and AIDS, 2018 ncbi.nlm.nih.gov*

Sharaf, M.F., Rashad, A.S., 2018. Socioeconomic inequalities in infant mortality in Egypt: *analyzing trends between 1995 and 2014. Soc. Indic. Res. 137, 1185–1199*