MODELLING THE RISK INDEX OF INFANT MORTALITY IN NIGERIA

By

Iseh, Matthew Joshua[1], Ph.D. & Udoh, Nse Sunday[2], Ph.D.

Department of Statistics, Akwa Ibom State University, Mkpat Enin, Nigeria &
Department of Statistics, University of Uyo, Uyo, Nigeria

ABSTRACT

This study models the determinants of infant mortality in Nigeria. Among other factors, this study examined some risk factors considered as basic determinants associated with infant mortality to include; types of residence, Wealth index, Sex of the Child, Mothers' level of education, parity, birth order, source of drinking water and toilet facilities. The determinants of 23666 households in Nigeria were extracted from the Nigeria Demographic and Health Survey (DHS) 2013. This study adopts the Logistic regression technique to analyze the determinants of infant mortality in Nigeria. Results obtained revealed that five determinants have significant influence on infant mortality, and this include; wealth index, mothers' level of education, sex of the child, parity, and birth order. In addition, the result further indicates that the odds of infant mortality were higher in households who were either poorer or poorest. However, it is noticeable in the result that as household wealth index increases, the odds of infant mortality decreases significantly.

KEYWORDS: *Determinants, Logistic regression, Odds ratio, Households*

1.0 Introduction

Infant mortality is a tragedy, which many parents are rendered childless. It is defined as the death of a child before its first birthday. United Nation International Children Emergency Fund (UNICEF) 2008 defined infant mortality as the death of a live born child between birth and a span of 12-month. Infant mortality is known to be one of the most sensitive and commonly used indicators of the socio-economic developments of a country.

Since infant is one of the age-groups in the population that depends heavily on the socio-economic conditions of their environment for survival, thus, the level of infant mortality

[1] Corresponding Author: Dr. Iseh Matthew J. Email: matthewiseh@aksu.edu.ng
ORCID ID: https://orcid.org/0000-0003-2696-7319

would present a measure of how well a society meets the needs of its people, Bicego and Ahmad (1996). Health is a state of human well-being, which was declared by the United Nation as a right. Thus, striving for improvement in health is a moral obligation for policy-makers at all levels of governance (National and International).

Infant mortality is a useful indicator on the Nation's health because it is often associated with health factor such as maternal health. Statistics from the Save Children Organization, an international non-profit group has revealed that Nigeria is one of the countries with high rate of infant mortality in Africa, Lawoyin (2007).

Nwaokoro et al (2015) examined the risk factors associated with infant mortality in Owerri Metropolis, Imo State, South Eastern Nigeria. Results obtained showed that pre-pregnancy factors, previous number of children, birth spacing, previous birth complications, Human Immune Deficiency Virus (HIV) infection, and malnutrition contributed to infant mortality. Other factors include; alcohol intake, level of mother's education, previous caesarian section, antenatal factors (low birth weight, gestational diabetes, failure to receive tetanus toxoid vaccine, congenital malformation, alcohol, smoking or staying near a smoker, malaria in mother, obesity, feeding habit of a pregnant women). Post-natal factors (over weight of a baby, place of delivery, birth attendant, preterm birth, length of labor, placenta abnormalities, caesarian section, failure to have a wellness baby check, jaundice) were observed to contribute significantly to infant mortality in Owerri Metropolis.

Anyamelel, Akanegbu, Assad and Ukawuilulu (2017) examined the zonal differences in the role wealth, education and religion play in child and infant mortality in Nigeria using 2003 and 2008 Nigeria demographic and health survey (DHS) pooled data. The study used logistic regression technique to obtain the odds ratio of which group has lower or higher odds of child mortality based on wealth quintile and on geographic location in various zones of Nigeria. The findings showed that education and wealth are significant factors in explaining the urban-rural differences in infant and child mortality rates in Nigeria. They also found that the risk of both infant and child mortality is higher in the North West and North East Zones of Nigeria than any other zones. In addition, the South West region has the lowest risk of both infant and child mortalities in Nigeria. There was no evidence of statistical significant difference in the risk of both infant and child mortalities between the urban and 'rural poorer and poorest wealth quintiles in Nigeria. The study established the differentials in infant and child mortality in the six geo-political zones of Nigeria. The findings also showed that there is a disparity in both infant and child mortalities between the urban and rural areas.

The determinants of infant mortality are not static, they vary with geographical locations, and there are various determinants, classified into four categories, which contribute to the infant mortality. This include;

(i) Socio-economic factors, which includes mother's level of education, occupation, residence, type of delivery, wealth index and medical care.

(ii) Environmental variables (factors) include source of water, toilet facilities, distance from home to the nearest health service and sources of energy.

(iii) Demographic factors are age of the mother at childbirth, birth order, previous birth interval breastfeeding, and sex of the child.

(iv) Cultural factors; e.g. Religion.

Despite all measures taken by government, UNICEF, WHO, health initiates in the country such as building dispensaries, hospital and health centers, infant's death are still increasing especially in the rural population. Many people have little or no knowledge of what could be the major causes of infant mortality. This study considers factors responsible for infant mortality in the Nigeria to include; type of residence, mother's level of education, source of drinking water, wealth index, birth order, and sex of the child, parity and toilet facilities. We therefore, seek to examine which among these risk factors is mostly associated with high infant mortality in Nigeria. Hence, this research would determine among other things which of the listed factors contribute significantly to infant mortality in Nigeria by using multiple logistic regression to model the effect of each determinant, and to fit a predictive reduced logistic regression model to the significant factors. Of a greater importance is to identify which of the determinants is of high risk to infant mortality.

1.1 Forms of Infant Mortality

According to Ouma, Bashar and Tuno (2014), the states of infant mortality are:

(1) Neonatal Mortality: This is the death of newborn occurring within 28 days post-partum. Neonatal death is often attributed to inadequate access to basic medical care during pregnancy and after delivery.
(2) Post-Neonatal Mortality: Is the death of babies between 29 and 365 days of life. The major contributions to post neonatal death are malnutrition, infectious disease and problem with the environment.

(3) Prenatal Mortality: This involves death of a child in the first week of life including stillbirths. It is at high level in developing countries especially those in sub-Saharan African.

1.2. Review of Empirical studies

There are socioeconomic, demographic, environmental and cultural factors, which contribute to infant mortality. Reducing infant mortality should be a top priority in the national public health objectives. Its relative effect of each factor differs from one population to another, as reported by Gwatkin et al (2000). It is of interest to study their association and their contribution in the population. Household socioeconomic status (Wealth index) is important for child survival. It determines the amount of resources (such as food, good sanitation and health care) that is available to the child, Millard (1994).

Curtis and Steele (1996) used data from Demographic and Health Survey (DHS) from Bolivia, Peru, Kenya and Tanzania to study neonatal mortality using logistic regression. They found that the level of maternal education was highly significant. It also showed that infant mortality was higher in poorer households than in urban areas. Brockerhoff (1993), Madise and Diamond. (l995), observed that Mothers' level of education can affect the child survival by influencing her choices and increasing/limiting her skills in health care practices related to contraception, nutrition, hygiene, preventative care and disease treatment. The relationship between mothers' education and child survival has received many attentions. It is one of the most frequently described social determinants of infant mortality in developing countries where the parent's education is weaker (mother's education in particular), there is high prevalence of infant mortality.

Machado and Hill (2005) and Smith and Evans (2009) established relationship between types of dwelling and infant mortality using Cox regression. The authors observed that households in most of the urban areas where social services are near to peoples' residence had a reduction in the odds of neonatal deaths compared to rural areas. Shehzad (2006) found that, in Pakistan, child illnesses such as diarrhea, acute respiratory infections and fever are affected by family size, housing and parental education.

Many studies have shown that infant mortality are influenced by numbers of demographic factors, such as maternal age, birth order, parity, breastfeeding, sex of the child, preceding birth interval and survival of preceding sibling(s), (Quamral, Islam and Hossain 2010, Elamin and Bhayan, 1999). The 2008 Nigeria demographic and health survey concluded that demographic factors contributed to the global discourse by examining the influence of community level characteristics (as opposed to individual level attribute) on infant mortality in Nigeria. Result indicated that risk of infant death were almost twofold higher for babies residing in the North East and North West region compared with babies in South west region of the country. Risk of death were lower for children of mothers attending prenatal care by a doctor

Environmental Health Determinants of infant survival and environmental condition have long been considered to have a significant influence on infant mortality. These include access to sanitation, source of drinking water, source of energy, type of dwelling and toilet facility. Studies conducted by Anderson, Romani, Philips and Vanzyl (2002), and Wichmann and Voyi (2006) which employed Cox regression have shown that infant mortality rates were double where the source of drinking water was other than piped water, and poor sanitation existed. Infant mortality increased three times more where materials other than block/bricks are used for housing and other sources of energy other than electricity were being used. Mahmood (2002) also found that families living in the household with pipe borne water connected in their houses have a significantly lower post neonatal mortality than those families, which depend on well for drinking water.

Baraki et.al. (2020) assessed the determinants of infant mortality based on Ethiopia Demographic Health Survey data, and observed that Sex of the child, multiple births, duration of pregnancy, and residence were significantly associated with infant mortality. The study also support earlier findings by several authors that risk of infant mortality has also shown differences across different regions.

2.0 METHODOLOGY

This section focuses on source of data used, and methods of analysis employed to achieve the set objectives.

2.1 Source of Data

The data for this study was extracted from the Nigeria Demographic and Health Survey data (NDHS) for 2013 at https//www.unicef.org/Nigeria/publications_8559.html. However it suffices to state that the NDHS data were based on nationally representative sample of women age l5-49 years and men aged 15-59 years who were selected using a stratified two-stage cluster sampling technique (NDHS, 2013). The survey was implemented by the national population commission (NPC). They observed covariates characteristics such as type of residence, Wealth index, sex of the child, parity, mother’s level of education, birth order, source of drinking water and toilet facilities. The dependent variables are mortality is either child is alive or dead. If the child is alive, it is coded 0, but if dead is coded 1. Summary of these variables are presented in Table 1.

Table 1: Categorization of infant mortality

	Determinants	Category
1	Type of residence	l = rural, 2 = urban
2	Wealth index	l = poorest, 2 = poorer, 3 = middle, 4 = richer, 5=richest
3	Sex of the child	1 = male, 2 = female
4	Parity	1=l, 2=2-3, 3 = 4 and above.
5	Mother's level of education	0=No formal education, l=Primary, 2=secondary, 3=Higher education
6	Birth order`	l=l, 2=2-3, 3=4 and above
7	Source of drinking water	l = Improve, 2=Not improve
8	Toilet facilities	l = Improve, 2=Not improve

2.2 Methods of Data Analysis

This section encompasses the different methods that will be applied in the analysis of the data with one or more independent variables.

2.2.1 Logistic Regression

Logistic regression also called logistic model or logit model is a statistical method for analyzing data set in which there is one or more independent variables and categorical dependent variable to estimates the probability of occurrence of an event.

The outcome is measured with a dichotomous variable, where there are two possible outcomes example yes or no, success or failure and the dependent variables is dichotomous which is code as 1 and 0 which represent the outcome of alive or dead in this work. There are two models of logistic regression:

2.2.1.1 Binary logistic regression

Is use when the dependent variables are dichotomous and independent variables are either continuous or categorical. Binary logistic regression model is given by:

$\log\frac{(\pi_i)}{1-\pi_i} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon_i$, $\left(\frac{\pi_i}{1-\pi_i}\right) \sim B = 0{,}1$

2.2.1.2 Multiple Binary Logistic Regressions

Let $X_1, X_2, \ldots X_n$, be the sets of explanatory variables and Y, the n predictors for a binary response Y Logistic regression model is based on probabilities associated with the outcome value of Y. We use $\pi(x)$ to represent probability that $Y = 1$ for the success and $1 - \pi(x)$ to represent probability that $Y = 0$. Hence, these probabilities are written in the following form:

$$\pi(x) = P(Y = 1|X_1, X_2, \ldots X_n)$$

and

$$1 - \pi(x) = P(Y = 0|X_1, X_2, \ldots X_n)$$

The model for log odds is

$$\text{L}ogit\ \pi[(x)] = \ln\frac{P(Y = 1|X_1, X_2, \ldots X_n)}{P(Y = 0|X_1, X_2, \ldots X_n)}$$

Which gives

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

Or

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \sum_{j=1}^{n}\beta_j X_j + \varepsilon$$
(1)

It is deduced by taking exponential function

$$\frac{\pi(x)}{1-\pi(x)} = \ell^{\ \beta_0+\sum_{j=1}^{n}\beta_j X_j+\varepsilon}$$
(2)

$$\pi(x) = \frac{\ell^{\ \beta_0+\sum_{j=1}^{n}\beta_j X_j+\varepsilon}}{1+\ell^{\ \beta_0+\sum_{j=1}^{n}\beta_j X_j+\varepsilon}}$$

$$\pi(x) = P(Y = 1|X_1, X_2, \dots X_n) = \frac{\ell^{\beta_0 + \sum_{j=1}^{n} \beta_j X_j + \varepsilon}}{1 + \ell^{\beta_0 + \sum_{j=1}^{n} \beta_j X_j + \varepsilon}} \quad ; \; 0 < \pi(x) < 1$$

(3)

β_0 is the constant; $\exp(\beta_i)$ tell how the odds of the response (y) increase as X_j increase by one unit.

2.2.1.3 Assumptions of Binary Logistic Regression Model

- Logistic regression requires the dependent variable to be discrete mostly dichotomous.
- Logistic regression estimates the probability of an event occurring P(Y=1). It is necessary to code the dependent variable accordingly, that is, the desire outcome should be coded to be 1.
- Logistic regression requires each observation to be independent to avoid multi-collinearity.
- Logistic regression does not require a linear relationship between the dependent and independent variables; it requires that the independent variables are linearly related to the odds of an event.
- Large sample sizes are required because maximum likelihood estimate is use to estimates the unknown required parameters.

2.2.2 Maximum Likelihood Estimation of Logistic Regression Parameters

The goal of logistic regression in Eq. 4 is to estimate the $N + 1$ unknown parameters, β which is $\beta_0, \beta_1, \; \beta_2 \dots \beta_n$. This is done by using maximum likelihood estimation which entails finding the set of parameters for which the probability of the observed data is optimal. The maximum likelihood estimation is derived from Bernoulli probability distribution of the dependent variable. For a set of observation in the data (x_i, y_i) the likelihood function (L) of the sample data as the product across all sampled cases of the probabilities, the likelihood function for $\pi(x_i), when\; y_i = 1, and\; 1 - \pi(x_i), when\; y_i = 0$ is given as follows: let $\phi\,(x_i)$ be the likelihood function for the observation (x_i, y_i).

Hence:

$$\phi\,(x_i) = \pi(x_i)^{y_i}\,[[1 - \pi(x_i)]^{1-y_i}\,]$$

(4)

Applying the likelihood function

$$L(\beta) = \prod_{i=1}^{n} \phi(x_i) = \prod_{i=1}^{n}[\pi(x_i)^{y_i}[1-\pi(x_i)]^{1-y_i}]$$

$$L(\beta) = \pi(x_i)^{\sum y_i}[1-\pi(x_i)]^{-\sum y_i}[1-\pi(x_i)]^{n}$$

$$L(\beta) = \left[\frac{\pi(x_i)}{1-\pi(x_i)}\right]^{\sum y_i}[1-\pi(x_i)]^{n} \qquad (5)$$

Substituting Eqs. (2) and (3) into Eq. (5) gives

$$\Rightarrow L(\beta) = \left(\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}\right)^{\sum_{i=1}^{n} y_i}\left(1-\frac{\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}}{1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}}\right)^{n} \qquad (6)$$

$$L(\beta) = \left(\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}\right)^{\sum_{i=1}^{n} y_i}\left(\frac{1}{1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}}\right)^{n}$$

$$L(\beta) = \left(\ell^{\beta_0\sum_{i=1}^{n} y_i+\sum_{i=1}^{n} y_i\sum_{j=1}^{n}\beta_j X_j}\right)\left(1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}\right)^{-n} \qquad (7)$$

In Eq. (3), β contains all the parameters β_0 , $\beta_1, \ldots$ β_n and $L(\beta)$ is the likelihood function β. The maximum likelihood estimators (MLE's) β_0 , $\beta_1, \ldots$ β_n is obtained by β which maximizes $L(\beta)$.

Taking the natural logarithm of the Eq. (7)

$$ln[L(\beta)] = ln\left[\left(\ell^{\beta_0\sum_{i=1}^{n} y_i+\sum_{i=1}^{n} y_i\sum_{j=1}^{n}\beta_j X_j}\right)\left(1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}\right)^{-n}\right] \qquad (8)$$

$$ln[L(\beta)] = \left(\beta_0\sum_{i=1}^{n} y_i+\sum_{i=1}^{n} y_i\sum_{j=1}^{n}\beta_j X_j\right) - nln\left(1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}\right) \qquad (9)$$

Finding the maximum likelihood estimates requires taking first derivative on the log likelihood function $L(\beta)$ with respect to the parameters:

$$\frac{\partial L(\beta)}{\partial L(\beta_0)} = \sum_{i=1}^{n} y_i - n\left(\frac{\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}}{1+\ell^{\beta_0+\sum_{j=1}^{n}\beta_j X_j}}\right) = 0$$

$$\sum_{i=1}^{n} y_i - n\pi(x_i) = 0$$

$\sum_{i=1}^{n}[y_i - \pi(x_i)] = 0$
(10)

with respect to β_j

$$\frac{\partial L(\beta)}{\partial L(\beta_j)} = \sum_{i=1}^{n} y_i \sum_{i=j}^{n} x_j - n \sum_{i=j}^{n} x_j \left(\frac{\ell^{\ \beta_0 + \sum_{j=1}^{n} \beta_j x_j}}{1 + \ell^{\ \beta_0 + \sum_{j=1}^{n} \beta_j x_j}} \right) = 0$$

$$\frac{\partial L(\beta)}{\partial L(\beta_j)} = \sum_{i=1}^{n} y_i \sum_{i=j}^{n} x_j - n \sum_{i=j}^{n} x_j \pi(x_i)$$

$$\frac{\partial L(\beta)}{\partial L(\beta_j)} = \sum_{i=j}^{n} x_j \left[\sum_{i=1}^{n} y_i - n\pi(x_i)\right]$$
(11)

The maximum likelihood estimate for $\hat{\beta}_0$ and $\hat{\beta}_j$ can be found by setting each of the $N + 1$ parameters of β in Eqs. (10) and (11) to zero and solve for each $\hat{\beta}_0$ and $\hat{\beta}_j$. This was carried out with the help of R- software, SPSS version 20.0

2.2.3 Multi-collinearity

Multi-collinearity exists when a regression model are moderately or highly correlated. When it exists any of the following pitfalls can be exacerbated;

- The hypothesis test for $\beta_j = 0$ may yield different conclusion depending on which predictors are in the model.
- The precision of the estimated regression coefficient decreases as more predictors are added to the model.
- The marginal contribution of any one predictor variable in reducing the error sum of squares depends on which other predictors already in the model.
- The estimated regression coefficient of any one variable depends on which other predictors are included in the model.

2.2.4 Detecting Multi-collinearity

Tolerance value/variance inflation factor (VIF) is used to detect the presence of multicollinearity, it quantifies how much the variance is inflated. Standard errors and variances of the estimated coefficients are inflated when multicollinearity exists. Multicollinearity is said to be present in the model if the tolerance values are low (less than 0.1) and the variance inflation factor is 10 and large (Katz 1999, Iseh et.al. 2022)

Let the variance inflation factor for the estimated coefficient β_j be denoted as VIF_j. For the model in which x_j is the only predictor;

$$y_i = \beta_0 + \beta_j X_{ij} + \varepsilon_i$$

The variance of the estimated coefficient b_j is

$$var(b_j)_{min} = \frac{\delta^2}{\sum_{i=1}^{n} \left(x_{ij} - \overline{x}_j\right)^2} \quad (12)$$

The subscript (min) denote the order of how small the variance can be.

For model with correlated predictors.

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_j X_{ik} + \cdots + \beta_{p-1} X_i + \varepsilon_i \quad (13)$$

If some of the predictors are correlated with predictor X_j , then the variance of b, is inflated. It is shown that variance of b is

$$var\left(b_j\right) = \frac{\delta^2}{\sum_{i=1}^{n} \left(x_{ij} - \overline{x}_j\right)^2} \times \frac{1}{1 - R_j^2} \quad (14)$$

R_j^2 is the R^2- values obtain by regressing the jth predictors on the remaining predictors. The greater the linear dependence among the predictor X_j and other predictors, the larger the R_j^2 value. Eq. (14) suggests that, the larger the R_j^2 value, the larger the variance of b_j.

To answer the question, ‘how large’? We take the ratio of the two variances.

$$\frac{var(b_j)}{var(b_j)min} = \frac{\frac{\delta^2}{\sum_{i=1}^{n}\left(x_{ij}-\overline{x}_j\right)^2} \times \frac{1}{1-R_j^2}}{\frac{\delta^2}{\sum_{i=1}^{n}\left(x_{ij}-\overline{x}_j\right)^2}}$$

$$VIF_j = \frac{1}{1-R_j^2}$$

(15)

Eq. (15), is the variance inflation factor which is reciprocal of tolerance value $1 - R_j^2$

2.2.5 Odds Ratio (OR)

An OR is a measure of association between an exposure and an outcome; it represents the odds that an outcome (e.g. disease or disorder) will occur given a particular exposure; compared to the odds of the outcome occurring in the absence of that exposure. OR is used to determine whether a particular exposure is a risk factor for a particular outcome, and compared the magnitude of various risk factors for that outcome.

When a logistic regression is analyzed, the regression coefficient b_j is the estimated increase in the log odds of the value of the exposure, in other words the exponential function of the regression coefficient e^{b_j} is the odds ratio associated with a one-unit increase in the exposure.

Mathematically,

$$\text{Odds Ratio (OR)} = \frac{\text{\# of exposed cases given number of unexposed cases}}{\text{\# of exposed non cases given number of unexposed non cases}}$$

2.2.6 Test of Hypotheses for Logistic Regression

$H_0: \beta_j = 0$ Vs $H_1: \beta_j \neq 0$

$Let\ \alpha = 0.05$

2.2.7 Decision Rule

If Wald Test $\geq x^2$ critical, H_0 is said to be significant at α = 0.05 level else H_0 is not significant at α = 0.05 level.

Odds Ratio with 95% confidence interval (CI): is used to assess the contribution of individual predictors. 95% confidence interval is used to estimate the precision of the OR.

A large CI indicates a low level of precision of the OR whereas a small CI indicates a higher precision of the OR. An approximate confidence interval for the population log odds ratio is; 95% CI for the

$\ln(OR) = \ln(OR) \pm 1.96 \times SE \ln(OR)$, taking the antilog, 95% CI for OR is $\ell^{\ln(OR)\ 1.96 \times SE(OR)}$

2.2.8 Wald Test Statistics

Wald statistic is used to test the significance of individual coefficient in the model. The Wald statistic is the ratio of the square of regression coefficient to the ratio of the square of standard error of the coefficient.

$$w_j = \left(\frac{coefficient\ of\ the\ parameter}{standard\ error\ of\ the\ cooefficient\ f\ the\ parameter}\right)^2$$

$$w_j = \left(\frac{B_j}{SE(B_j)}\right)^2$$

Each Wald statistic is compared with a x^2 distribution with degree of freedom. The degree of freedom in Logistic Regression is the number of predictors added to the model.

3.0 DATA ANALYSIS, AND DISCUSSION OF RESULTS

3.1 Data Presentation

Data collected for this work comprises of type of residence, wealth index, sex of the child, parity, mother's level of education, birth order, source of water, and toilet facilities. The dependent variable (mortality) is either baby is alive or dead. If the baby is alive, it is coded (0) but if baby is dead, it is coded (1). The data is comprised of 23666 household of the six geographical zones in Nigeria.

3.2 Data Analysis

Data obtained were analyzed using multiple logistic regression and Statistical Package for Social Science (SPSS) version 20.0 and they are presented in Tables.

Table 2: Tolerance values and variance inflation factor for each variable

Characteristics (factor)	Tolerance value	Variance inflation factor
Type of residence	0.635	1.574
Wealth index	0.606	1.650
Sex of the child	0.999	1.001
Mother's level of education	0.749	1.259
Birth order	0.744	1.344
Parity	0.699	1.431
Source of drinking water	0.963	1.038
Toilet facilities	0.957	1.045

Result in Table 2 shows that all the variables have tolerance value greater than 0.1 and VIF less than 10. This means that there is no evidence of multicollinearity among the independent variables.

Table 3: Multiple Logistic Regression showing the Influence of some factors on infant mortality with 95% confidence interval

Factors	B	S.E	Wald	Odds ratio (OR) 95% C.I	df	x^2-critic
Type of residence			0.001		1	3.84
Urban	-	-	-	1.00(reference)		
Rural	0.002	0.066	0.001	1.002(0.880 – 1.140)		
Wealth index			28.907		4	9.49
Poorest	-	-	-	1.00(reference)		
Poorer	-0.007	0.097	0.005	0.993(0.821 - 1.201)		
Middle	-0.214	0.102	4.402	0.807(0.661 – 0.986)		
Richer	-0.342	0.105	10.609	0.710(0.578 – 0.873)		
Richest	-0.399	0.104	14..719	0.671(0.547 – 0.823)		
Sex of the child			8.873		1	3..84
Male	-	-	-	1.00(reference)		
Female	-0.140	0.047	8.873	0.869(0.793 – 0.953		

Mother's level of education			10.072		3	7.82
No formal	-	-	-	1.00(reference)		
Primary	-0.721	0.262	7.573	0.486(0.291 – 0.813)		
Secondary	-0.705	0.268	6.921	0.494(0.292 – 0.835)		
Higher	-0.827	0.266	9.666	0.437(0.260 – 0.737)		
Birth order			75.816		2	5.99
1	-0.542	0.065	69.530	0.582(0.512 – 0.661)		
2 – 3	-0.109	0.057	3.664	0.897(0.802 – 1.003)		
4 and above	-	-	-	1.00(reference)		
Parity			205.663		2	5.99
1	-	-	-	1.00(reference)		
2 – 3	1.088	0.132	67.938	2.968(2.292 – 3.843)		
4 and above	0.948	0.073	168.644	2.581(2.237 – 2.977)		
Source of drinking water			0.250		1	3.84
Improved	-	-	-	1.00(reference)		
Not improved	0.033	0.066	0.250	1.034(0.908 – 1.176)		
Toilet facilities			0.003		1	3.84
Improved	-	-	-	1.00(reference)		
Not improved	0.007	0.138	0.003	1.007(0.768 – 1.320)		

3. 3. Discussion of Results

3.3.1. Influence of Type of Residence on Infant Mortality

H_{01}: There is no significant influence of type of residence on infant mortality. The result of the multiple logistic regression summarized in Table 3 revealed that there is no significant influence of type of residence on infant mortality in Nigeria with a Wald Statistic = 0.001 $< x^2$-critic =3.84. The result also shows that the odds of infant mortality in the rural is not significantly different from that of the urban (OR = 1.002, C.I = 0.880 — 1.140). The null hypothesis is retained that there is no significant influence of type of residence on infant mortality.

3.3.2. Influence of type of Wealth Index on Infant Mortality

H_{02}: There is no significant influence of wealth index on infant mortality.

As shown in Table 3, there is a significant influence of wealth index on infant mortality in Nigeria with a Wald statistic = 28.907 $> x^2$-critic. = 9.49. Wealth index which was grouped into five categories; poorest, poorer, middle, richer and richest, with poorest group as the reference category, shows that the odds of infant mortality decreases as the household wealth index increases. The result also reveals that the odds ratio of infant mortality among those within the categories; middle (OR = 0.807, C.I = 0.661 — 0.986), richer (OR = 0.710, C.I = 0.579 — 0.872), richest (OR = 0.671, C.I = 0.578 — 0.873) were significantly less than that obtained for households who were poorest. Result also shows that there is no significant difference in the odds of infant mortality between households who were poorest and poorer (OR = 0.993, C.I = 0.821 — 1.201). Therefore, the null hypothesis is rejected, affirming that there is significant influence of wealth index on infant mortality in Nigeria. The odds ratio of Binary Multiple logistic regression has a control pattern which also revealed that the poorer category experienced higher infant mortality, compared to the reference category (the poorest), followed by middle category, richer, whereas the richest category experienced low infant mortality with. This findings on wealth index agrees with that of Anyamelel, Akanegbuz, Assad and Ukwwuilulu (2017) which examined the role of wealth quintiles influence on infant mortality using logistic regression technique, the result shows that wealth index has significant influence on infant mortality.

3.3.3. Influence of Sex of the Child on Infant Mortality

H_{03}: There is no significant influence of sex of the child on infant mortality.

The result displayed in Table 3 shows that there is a significant influence of sex of the child on infant mortality in Nigeria. The risk factor was categorized into two groups; the male and the female, of which male were the reference category. The result shows that the odds of infant mortality is lower in female compared to their male counterparts, with odds ratio 0.869 (0.793-0.953) with a Wald test = 8.873> x^2 critical = 3.84. Hence, null hypothesis is rejected, that there is a significant influence of sex of the child on infant mortality in Nigeria.

3.3.4 Influence of Mother's Level of Education on Infant Mortality

H_{04}: There is no significant influence of mother's level of education on infant mortality.

The result in Table 3 shows significant influence of mother's level of education on infant mortality in Nigeria with a Wald test = 10.072 $> x^2$ critical = 5.99. The influence of mothers' level of education factor was group into four categories; no formal education, primary, secondary and higher education, of which mothers without formal education were the reference category. The result also reveals that the odds of infant mortality among those whose mother had primary education (OR =0.486, C.I =0.291—0.8l3), Secondary education (OR = 0.494, C.I =0.292—0.835) and Higher education (OR =0.437, C.I =0.260—0.737) were significantly less than that of infants born to mothers who have no formal education. Therefore, there is a significant influence of mothers' level of education on infant mortality in Nigeria. The findings is also in line with Sharifzadeh et.al. (2008) who analyzed the infant and child mortality using descriptive statistic, ANOVA and regression analysis, which resulted in high significant influence of mothers' level of education on infant and child mortality.

3.3.5. Influence of Birth Order on Infant Mortality

H_{05} There is no significant influence of birth order on infant mortality.

The result displayed in Table 3 reveals that there is a significant influence of birth order of infants on their mortality with a Wald statistic =75.816 $> x^2$-critic. = 5.99. Birth order which was also considered as a determinant factor was grouped into three categories; birth order (1), birth order (2-3) and birth order (4 and above), of which birth order 4 and above were the reference category. The result shows that household with birth order (1) had lower odds of infant mortality than household with birth order (2-3), and 4 and above with odds ratio 0.582 (0.512- 0.661). It also shows that in household with birth order (2-3), the odds of infant mortality were lower than those with birth order 4 and above with odds ratio 0.897 (0.802-1.003). The result of this study also reveals that there is a significant influence in birth order of infant mortality. It means that as the birth order increases, the odds of likelihood of infant mortality increases. This findings agree with that by Kittur (2014) which established that birth order has high significant influence on infant mortality in urban Kenya.

3.3.6. Influence of Parity on Infant Mortality

H_{06}: There is no significant influence of parity on infant mortality.

The result displayed in Table 3 reveals that there is a significant influence of parity on infant mortality with a Wald statistic = 205.663 $> x^2$-critic. =5.99. Parity was categorized into three groups; category one, household with parity (1), category two, household with parity (2-3), and category three, household with parity (4 and above), of which parity (1) was the reference category. The result shows the odds of infant mortality is 3 times higher with those in parity (2-3), parity 4 and above than those in parity (1). The OR for parity (2-3) 2.968 (2.292—3.843) and that of parity 4 and above 2.581 (2.237 -2.977) were significantly higher than that of household with parity 1. Again, this study shows significant influence of parity on infant mortality.

3.3.7. Influence of Source of Drinking Water on Infant Mortality

H_{07}: There is no significant influence of source of drinking water on infant mortality.

The result displayed in Table 3 reveals that there is no significant influence of source of drinking water on infant mortality with a Wald statistic = 0.250 $< x^2$-critic. = 3.84. The result also reveals that the odds of infant mortality between households who have improved and non-improved source of drinking Water was not significantly different (OR = 1.034, C.I = 0.908 — 1.176). Therefore, the null hypothesis is retained. Hence there is no significant influence of source of drinking water on infant mortality.

3.3.8 Influence of Toilet Facilities on Infant Mortality

H_{08}: There is no significant influence of toilet facilities on infant mortality.

The result displayed in Table 3 reveals that there is no significant influence of toilet facilities on infant mortality with a Wald statistic =0.003 $< x^2$-critic. = 3.84. The result also shows that the odds of infant mortality between households who had improved and non-improved toilet facilities was not significantly different (OR = 1.007, C.I = 0.768 — 1.320). Therefore, the null hypothesis is retained. Hence, there is no significant influence of toilet facilities on infant mortality in Nigeria.

3.3.9 Fitting the General Logistic Regression based on the Risk Factors Considered

The general model for the multiple logistic regressions for the eight-predictor variables; recall Eq. (3)

$$\pi(x) = \frac{\ell^{\,\beta_0 + \sum_{j=1}^{n} \beta_j X_j}}{1 + \ell^{\,\beta_0 + \sum_{j=1}^{n} \beta_j X_j}}$$

$$Let\ y = \pi(x)$$

$$y = \frac{\ell^{\,\beta_0 + \sum_{j=1}^{8} \beta_j X_j}}{1 + \ell^{\,\beta_0 + \sum_{j=1}^{8} \beta_j X_j}} \tag{16}$$

Based on the result presented in Table 3: the estimates of the parameters of multiple logistic regressions; given in Table 4 as follows:

Table 4: Estimation of parameters of the Full Multiple Binary Logistic Regression.

Determinants	$\boldsymbol{\beta}$
Type of residence (X_1)	-0.015
Wealth index(X_2)	0.115
Sex of the child (X_3)	0.144
Mother's level of education(X_4)	-0.001
Birth order(X_5)	0.255
Parity(X_6)	0.730
Source of drinking water (X_7)	-0.029
Toilet facilities(X_8)	-0.021
Constant	2.467

From Table 4, Estimates of the β_j are as follows;

β_0 = 2.267, β_1= -0.015, β_2= 0.115, β_3= 0144, β_4 = -0 00 1, β_5= 0.255, β_6 =0.730, β_7 = -0.029, β_8 = -0.021.

Substituting for $\beta_1, \beta_2, \dots, \beta_8$, in Eq. (16)

$$y = \frac{exp(2.467 - 0.015X_1 + 0.0115X_2 + 0.144X_3 - 0.001X_4 + 0.255X_5 + 0.730X_6 - 0.029X_7 -}{1 + exp(2.467 - 0.015X_1 + 0.0115X_2 + 0.144X_3 - 0.001X_4 + 0.255X_5 + 0.730X_6 - 0.029X_7}$$

3.3.10 Fitting the Reduced Logistic Regression Model Based on the Risk Factors.

From Table 3, is evident that the risk factors like type of residence (X_1), source of drinking water (X_7) and toilet facilities (X_8) were not significant based on the test statistic, and hence have no contribution(s) in the model. Hence, a reduced model estimates only on determinants that contributed significantly to the model is given in Table 5. This agrees with the result in the literature as observed by Ekong et. al. (2021).

Table 5: Estimation of the Parameters of the Reduced Multiple Logistic Regression

Factors	β
Wealth index (X_2)	0.043
Sex of the child (X_3)	0.123

Mother's level of education (X_4)	-0.021
Birth order (X_5)	0.213
Parity (X_6)	0.735
Constant	2.633

From Table 5, Estimates of the β_j are as follows;

β_0 = 2.633, β_2= 0.043, β_3= 0.123, β_4 = -0.021, β_5= 0.213, β_6 =0.735.

From Eq. (16), the reduced logistic regression model is;

$$y = \frac{exp(2.467 + 0.043X_2 + 0.123X_3 - 0.021X_4 + 0.213X_5 + 0.735X_6)}{1 + exp(2.467 + 0.043X_2 + 0.123X_3 - 0.021X_4 + 0.213X_5 + 0.735X_6)}$$

4. Conclusion

This study has identified the determinants of infant mortality in Nigeria. The result of the analysis has shown that five risk factors which include; wealth index, sex of the child, mothers' level of education, birth order and parity have significant influence on infant mortality. Therefore, this study concludes that these factors are the major determinants of infant mortality in Nigeria, and parity has the highest risk factor. This result has provided baseline information for government and health sectors to take proactive measures in curtailing infant mortality. The study has also provided a wider perspective to women on which of the observed risk factors are highly responsible for infant mortality, and the need to take appropriate care during their neonatal, post-neonatal and prenatal period. Above all, the study has provided information for planning and policies making on maternal health care in Nigeria. Hence, intervention programs should focus on these factors to help salvage the panic incidence of infant mortality in Nigeria.

REFERENCES

Anderson, B.A., Romani, J.H., Phillips, H. E. and Van-Zyl, J.A. (2002). Environment, access to health care and other factors affecting infant and child survival among the African and colored populations of South Africa. *Population and Environment*, 23(4): 34-78.

Anyamele, O.D, Akanegbu, B.N; Assad, J.C, and Ukawuilulu, J.O. (2017). Differentials in infant and child mortality in Nigeria: Evidence from pooled 2003 and 2008 DHS Data. *Advanced in Management and Applied Economics*, 7(6): 73-96.

Baraki, A. G., Akalu, T. Y., Wolde, H. F. Ayenew Molla Lakew, A. M., and Gonete, K. A. (2020). Factors affecting infant mortality in the general population: evidence from the 2016 Ethiopian demographic and health survey (EDHS); a multilevel analysis. *BMC Pregnancy and Childbirth* 20:299, https://doi.org/10.1186/s12884-020-03002-x.

Bicego, G. and Ahmad, O. B. (Eds.) (1996). Infant and child mortality: demographic and health surveys, Comparative studies No. 20. Macro International Inc., Calverton Maryland. 786pp..

Brockerhoff, M. (1993). Child survival in big cities. New York: Oxford University press

Curtis, S. L. and Steele, K. (1996). Assessment of the quality of data use for directs estimation of infant and child mortality in DHS-II Surveys, DHS Occasional papers No. 3. Macro International Inc., Calverton Maryland.

Elamin, M. A. and Bhuyan, K. C. (1999). Differential fertility in northeastern Libya. *The Journal of Family Welfare*, 45(1): 12 —22.

Ekong N., Moffat I., Usoro A., and Iseh M. (2021). A comparative study study of the impact of dummy variables on regression coefficients and canonical correlation indices: an empirical perspective. *International Journal of Analysis and application*, Vol. 19(4): 576-586.

Gwatkin, D. R., Rutstein, K., Johnson, R., Pande, S. and Waostafe, A. (2000). Socioeconomic differences in Health, nutrition and population in Zambia, HNP/poverty Thematic Group. The World Bank. Washington DC. 89pp.

Iseh, M., Ekong, N., Usoro, A., and Ukpe, I. (2022). Juxtaposing Vertically Transmitted Infections (VTIs) and the Spread of HIV/AIDS in a Typically Infection Prevalent Region in Nigeria. *Journal of the Society of Physical Sciences*, DOI:10,46481/jnsps.2022.418. Vol. 4(2022) 99-104.

Katz, M.H. (1999). Multivariable analysis: A practical guide for clinicians. Cambridge: Cambridge
University Press. '

Kittur, I.(2014). Factors influencing infant mortality in urban Kenya. A Master of Art Degree Thesis,
University of Nairobi, Kenya.

Lawoyin, T.O.(2007). Infant and maternal deaths in rural southwest Nigeria: a prospective study. *African*
Journal of Medical Science, 36(3): 235-241.

Machado, C. J. and Hill, K. (2005). Maternal, neonatal and community factors influencing neonatal
mortality in Brazil. *Journal of Biosocial Science*, (37): 193-208. "

Madise, N. J. and Diamond, I. D. (1995). Determinants of infant mortality in Malawi: An analysis to control
for death clustering within families. *Journal Biosocial Science*, 27: 95 - 106.

Mahmood, M.A. (2002). Determinants of neonatal and post-neonatal mortality in Pakistan. *The Pakistan*
Development Review 41(4): 723-744.

Millard, A. (1994). A casual model of high rates of child mortality. *Social Science and Medicine Journal,*
38: 253-268

Nigerian Demographic and Health survey Data (2013).

Nwaokoro, J .C, Ibfi-Sally, N.O, Ihenachor, C. A, Emerole, C.O, Nwufo, R.C., Ebirriekwe, C. and Onwuliri,
V. A. (2015). Risk factors associated with infant mortality in Owerri Metropolis, Imo, State. *Science Journal of Public Health*, 3 (5): 64-71.

Ouma, B. K. and Tuno, N. (2014). Seasonal abundance of anopheles mosquitoes and their association
with meteorological factors and malaria incidence in Bangladesh. *Parasite Vectors*, 7: 56-74.

Quamrul, C.H. Rafiqul, I. and Hossain, K. (2010). Socio-economic determinants of neonatal, post neonatal,
infant and child mortality. *International Journal of Socialist and Anthropology*, 2(6): 18-125.

Sharifzadeh, G.R, Namakin, K, and Mehrjoofard, H. (2008). An epidemiological study on infant mortality
and factors affecting it in rural areas of Birjand, Iran. *Iran Journal of Pediatrics*, 18 (4): 335-342.

Shehzad, S. (2006). The determinants of child health in Pakistan: An economic analysis. *Social Indicators*
Research, 8(3): 531-556.

Smith, E.G. (2009). Maternal schooling and child mortality in Nigeria: The importance of the actual
curriculum. Retrieved on February 18, 2018 from Princeton. edu/download. aspx? Submission ID =1 00377.

UNICEF, (April 2008). State of the world's children. Retrieved on January 12, 2018 from http//www.uniceforgjpublications/index l8108.htm.

Wichmann, J. and K. Vovi. (2006). Influence of cooking and heating fuel use on 1- 59 Month old Mortality
in South Africa. *Maternal Child Health Journal*, 10:553-561.

World Health Organization (August 1998). Fact sheet No. 94 World Health Organization. Retrieved on
January 25, 2018 from http//WWW.unitednation/childmortality.htm.