Gadau Journal of Pure and Allied Sciences
Gadau J Pure Alli Sci, 2(2): 156-161 (2023)
ISSN: 2955-1722
DOI: https://doi.org/10.54117/gjpas.v2i2.73

Modelling Child Survival at Birth Data Using Logistic Regression Model: A Case Study of Yobe State Specialist Hospital Damaturu, Nigeria

Umar Yusuf Madaki[1], Muhammad Ahmad[1] and Ishaq Abdullahi Baba[2]

[1]Department of Mathematics and Statistics, Faculty of Science, Yobe State University, Damaturu- Nigeria.

[2]Department of Mathematical Sciences, Faculty of Science, Taraba State University, Jalingo- Nigeria.

**Correspondence: uymadaki84@gmail.com; +2348035355649*

Abstract

This research examines factors influencing the survival of child at birth in Yobe State Specialist Hospital using a well-known logistics regression model. A total of 150 data points were collected through transcription from the maternity record of the Hospital to test significant factors that affect survival of child at birth. Analysis of logistic regression model was applied to the data using the Generalized Linear Model (GLM) package in R programming software. The results indicate that Type of delivery and weight at birth have the most significant influence on the probability of child survival at birth at 0.001 level of significance. Correlation analysis results show that all the five variables (age, parity, apgar, birth weight, and type of delivery) have a weak relationship, which implies that there is no multicollinearity in the data. Therefore, these results may help policy makers and health personnel to educate pregnant women on the effect of overweight baby at birth to reduce the incidence of deaths at birth. This study recommend that pregnant women should be educated about the effect of baby weight in a worm as that may increases the chance of caesarean section which in turn may affect the likelihood of child survival at birth. Further studies are suggested to consider factors like educational level, income of level of the family, antennal status, and blood pressure. A more advance community-based survey is recommended since not pregnant women attain formal health care facilities for antennal and postnatal services, which may expose more factors that influence child survival at birth.

Article History
Received: 25/03/2023
Accepted: 26/07/2023
Published: 25/08/2023

Keywords:
Child survival,
Child at birth,
Damaturu,
Modelling,
Logistic regression.

How to cite this paper: Madaki, U.Y., Ahmad, M., and Baba. I. A. (2023). Modelling Child Survival at Birth Data Using Logistic Regression Model: A Case Study of Yobe State Specialist Hospital Damaturu, Nigeria. *Gadau J Pure Alli Sci, 2(2): 156-161*. https://doi.org/10.54117/gjpas.v2i2.73.

1.0 Introduction

Giving birth is one of the most extra-ordinary experience in women life. The chances of survival of newborn baby has become an issue of concern both in Nigeria and globally. It has been noted that 2.4 million newborns died worldwide (Islam and Biswas, 2021). The prevalence of death at birth varies broadly depending on the geographical region. Sub-Sahara Africa has the highest neonatal deaths rate globally with 27 deaths per 1000 live births followed by central and south Asia with 23 deaths per 1000 live births (Emmanuel *et al*., 2021). The prevalence of neonatal in United State of America and Europe is 3.4 and 3.6 deaths per 1000 live births. Among other countries in Africa, Nigeria is rated as one of the countries with highest number of neonatal deaths as of 2020 when the neonatal death rate was recorded as 35.5 deaths per 1000 live births. This shows a rapid reduction from 65 deaths per 1000 live births in 1971 (Emmanuel *et al*., 2021). Recently, Ezeh *et al*. (2022) noted that Nigeria

Journal of the Faculty of Science, Bauchi State University Gadau, Nigeria
Phytochemistry and Natural Product

experience 1.1% increase in neonatal deaths between 2013 and 2018, signifying a setback to the former reported improvement in neonatal deaths in Nigeria. However, child survival at birth is determined right from the day one of the conception of the baby in the mother womb, the process of delivery, the expertise, the needed facilities available and the environment are factors that contribute to the neonatal deaths. That is child survival is primarily determined by the social and economic status of the child family essentially expressed by two indicators- maternal educations and some index of economic circumstances of the household. Some scholars investigated factors such as traditional belief, lacked adequate knowledge, access to the health care facilities and educational level of the husband or wife or both of them on the neonatal deaths (Emmanuel *et al*., 2021, Islam and Biswas, 2021, Esan *et al*., 2022, Madaki et al. 2023a, Madaki et al 2023b). The most important factors affecting mortality in childbirth are adequate nutrition and access to quality medical care. Medical care in this context does not refer specifically to treatment in hospital, but simply the presence of an attendant with midwifery skills. Furthermore, factors that affect childbirth negatively may include prematurity, high blood pressure, diabetes, and previous caesarean section.

One of the most dangerous risks to the fetus is that of premature birth and its associated low neonatal weight. The more premature (or underweight) a baby is the greater the risks for neonatal death and for pulmonary, respiratory, neurological, or other sequelae. About 12% of all infants born in the United States are born prematurely. In the past 25 years, medical technology has greatly improved the chances of survival of premature infants in industrial nations (Esan *et al*., 2022). In the 1950s and 1960s, approximately half of all low-birth-weight babies in the US died. Today, more than 90% survive. The first hours of life for premises are critical, especially the very first hour of life (Esan *et al*., 2022). Rapid access to a Neonatal intensive care unit is of paramount importance. Age limitation seems not to constitute impediment or barrier to pregnancy. In November 2004 Aleta. St. James, a 56-year-old single mother gave birth to twins conceived through in vitro fertilization. In 2005, a 67-year-old Romanian woman gave birth by caesarean to one surviving twin (Islam and Biswas, 2021). Based on the forgoing, several research works have been done regarding factors that influence the survival of child at birth. Ozuma and Nwogu-Ikojo (2008) discussed environmental factors of child mortality based on the principal component analysis were household environmental variable show significant impact on the mortality. Boerma and Bicego (1992) found that there is no consistent association between the biomedical and behavioural factors on the birth interval and child survival. They also noted that prenatal factors are more consequence than postnatal. Ronsmans (1996) examined the relationship between birth spacing and child mortality using the logistic and cox proportional hazard regression model were longitudinal data of 4852 born children between 1983 and 1989 is used. Maitra and Pal (2008) studied birth, spacing fertility and child survival using correlated hazard model. In Nigeria, studies on factors that contribute to neonatal death is revealing see for example Adetola *et al*. (2011), Akinyemi *et al.* (2015), and Audu *et al.* (2021). However, a small number of these studies focused on Northwestern parts of Nigeria. Furthermore, despite the high rate of neonatal deaths in Northeast, studies evaluating factors influencing neonatal deaths are limited in the literature. More studies focusing on neonatal deaths in northeast is necessary to average regional bias related to neonatal deaths. Therefore, the main purpose of this paper is to evaluate the risk factors associated with survival of child at birth using a logistic regression model. We will also conduct correlation test to investigate the degree of relationship among the risk factor relative to the survival of child at birth. In the remaining part of this paper, section 2 and 3 deliberated on the methodology and presentation data analysis and results. The last section gives the conclusion of the study.

2.0 Materials and Methods

Data analysis is a process of collecting, cleansing, transforming, and modelling data with the objective of gaining useful information or making a decision. Modeling procedure allows estimation of regression coefficients and selection of the significant and insignificant factors that contribute to the response variable (status of the child). Regression Analysis is a statistical technique used to investigate the relationship between the dependent and independent variables. Its main objective is to predict or estimate the value of the unknown regression parameters. There are numerous types of regression techniques depending on the type of data at hand. The method used to model child survival at birth is logistic regression. This type of regression is used when the dependent variable is binary-for example child status at birth (alive or death). Our main objective in this paper is to model the probability of child survival at birth using the logistic regression model. The logistic regression is usually used for three purposes: to predict the probability of the outcome variable, to categorize prediction, and to access the risk or odds associated with model predictors. The logistic regression Model formulation was well deliberated by Sperandei (2014), Wilson and Lorenz (2015), Boateng and Abaye, (2019). As mentioned earlier, the status of child at birth is specified to be either alive or death, with this binary response, the logistic regression model is the probability of an outcome based on the individual status. Consider a collection of p independent variables represented by the vector $x' = (x_1, x_2, \ldots, x_p)$. Let the conditional probability that the outcome present is represented by $p(y=1 \mid x) = \pi(x).$ The logic of the multiple logistic regression model is given by:

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \tag{1}$$

Where π represents the probability of an event for example child survival at birth (death or alive), $g(x) = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p$ with β_i representing the regression coefficient and x_i are the predictor variables. Fitting the model in equation (1) requires that we calculate estimate of the vector of regression parameters $\beta' = (\beta_0, \beta_1, \ldots, \beta_p).$ In this case, maximum likelihood procedure is applied for the estimation purposes. In the application of the likelihood procedure, there will be p+1 likelihood equations that can be achieved by differentiating the log likelihood function with respect to the p+1 regression coefficients. Thus, the likelihood equations can be expressed as:

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

$$\sum_{i=1}^{n} [y_i - \pi(x)] x_{ij} = 0, \quad j = 1,2,\ldots,p.$$

The solution of the log likelihood can be achieved using any available software such as the SAS or SPSS or Matlab or Python or R statistical software. In this paper the GLM package from the R statistical software is applied to fit the logistic regression model. The R is a free programming and application software introduced by Robert Gentleman and Ross Ihaka in 1990.

3.0 Data Analysis and Results

This study was designed to investigate the survival of newborn babies in Yobe State specialist Hospital Damaturu. The data were collected through transcription from the maternity records in the maternity unit of the Hospital. A total of 150 were sampled from the maternity records where individual file was used as data collection instrument. Note that all predictors are supposed to affect survival of child at birth (y). The variables associated with survival of child at birth includes mother age (x_1), parity (x_2), Apgar (x_3), Birth weight (x_4), Type of delivery (x_5). The main objective of this study is to develop a logistic regression model to predict the probability of survival of a child at birth. In addition, correlation analysis is conducted to measure the degree of relationship between the variables. In this application, the logistic regression model can be expressed as:

$$g(x) = y = x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4 + x_5\beta_5 + \beta_0$$

After applying the data to the model using a logistic regression R software package with Age, Parity, Apgar, Birth weight, Type of delivery as independent variables and child status as dependent variables, we obtained the output in Table 1. The maximum likelihood estimates of $\beta_0, \beta_1, \ldots, \beta_5$ are thus seen to be $\beta_0 = 0.498, \beta_1 = 0.004, \beta_2 = -0.012, \beta_3 = 0.143, \beta_4 = 0.151,$ and $\beta_5 = -0.175$. The fitted model is given by equation.

$$\hat{\pi}(x) = \frac{e^{0.498+0.004x_1-0.012x_2+0.143x_3+0.151x_4-0.175x_5}}{1+e^{0.498+0.004x_1-0.012x_2+0.143x_3+0.151x_4-0.175x_5}}$$

And the estimated logit function $g(x),$ is defined by the equation.

$$\hat{g}(x) = 0.498 + 0.004x_1 - 0.012x_2 + 0.143x_3 + 0.151x_4 - 0.175x_5$$

Table 1 display the log likelihood values computed based on the estimates of coefficients $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_5$. The second column show the standard errors of the estimated coefficients; the third column contain the p-values. Table 2 display the descriptive statistics of the data set. The descriptive statistics show the quantitative characteristics of the sample data points. Table 3 summaries the results of the correlation criteria to show the relationship that exist between the variables. The result in Table indicate that two independent variables (Birth weight and Type of delivery) are highly significant to the response variable (child status, death or alive) at 0.001 significant level. This means that increase in birth weight and Type of delivery are associated with child survival at birth. Even though the descriptive statistics in Table show small proportion of the death and delivery under Caesarean section.

Our model is able capture the associativity between the deaths as a result of delivery under Caesarean section. Based on the Table 3, we observed that the value of correlation coefficient of -0.04 indicate that there is negative weak relationship between the mother age and Apgar. Similarly, correlation coefficient of -0.09 indicates a weak relationship between the Parity and Type of delivery. Furthermore, correlation of 0.57 indicate a partially strong relationship between the mother's age and Parity. Since all the correlation values associated to the five independent variables under study has weak relationship, we can say that there is no problem of multicollinearity in the data used.

Hypothesis

$H_0 : \beta_i = 0$ **(Explanatory variables do not contribute significantly to child survival at birth)**

$H_1 : \beta_i \neq 0$ **(Explanatory variables contribute significantly to child survival at birth)**

Level of significance

α **= 0.05**

Table 1 Estimates of fitting Logistic Regression model

Variable	Coefficient	Std. Error	t-value	P>|t|
Constant	-2.788	4.129	-0.675	0.499
Age	0.094	0.131	0.719	0.472
Parity	-0.183	0.386	-0.476	0.633
Apgar	0.912	2.406	0.379	0.704
Birth weight	2.471	0.907	2.721	0.006
Type of delivery	-2.735	0.994	-2.752	0.005

Table 2: Descriptive statistics

Variable	Description	Mean	Min	Max	St. Deviation
Mothers age	continuous	29.79	0.00	42.00	5.27
Birth weight	continuous	1.71	0.00	6.00	1.55
Parity	continuous	0.73	0.00	1.00	0.20
Apgar score	continuous	2.92	0.00	4.10	0.61
		Frequency	Proportion		
Status of Child					
Y1	Death	8	0.05		
Y2	Alive	142	0.95		
Type of delivery					
1	Normal	134	0.89		
2	Caesarean section	16	0.11		

3.1 Testing Relationship between Variables in the Model

We use the correlation matrix or correlation table to show the relationships that exist between the variables. The variables in question are status of a child (Dead or Alive), mother's age, parity, APGAR score, birth weight and type of delivery.

Table 3 Correlation values

	Status of the child	Mother Age	Parity	Apgar score	Birth weight	Type of delivery
Status of the child	1.00	0.08	0.03	0.34	0.47	-0.34

• Mother Age	• 0.08	• 1.00	• 0.57	• -0.04	• 0.11	• 0.02
• Parity	• 0.03	• 0.57	• 1.00	• 0.03	• 0.09	• -0.09
• Apgar score	• 0.34	• -0.04	• 0.03	• 1.00	• 0.33	• -0.31
• Birth weight	• 0.47	• 0.11	• 0.09	• 0.33	• 1.00	• -0.08
• Type of delivery	• -0.34	• 0.02	• -0.09	• -0.31	• 0.08	• 1.00

3.2 Discussions

From table 1 and 2 the result we can see that the average mother's age is 29.79, with a standard deviation of 5.279, then the average of parity is 1.713 with a standard deviation of 1.551, the average APGAR score is 0.7269 with a standard deviation 0.200, the average birth weight is 2.92 with a standard deviation 0.606 and lastly the mean of type of delivery 1.113 with a standard deviation 0.357. Based on the above table 3, we observed that the value of the correlation coefficient -0.04, indicates that there is a weak negative association between Mother's Age and APGAR Score and the relationship is significant, also correlation coefficient -0.09 indicates that there is weak negative association between Parity and type of Delivery, correlation coefficient -0.31 indicates that there is a weak negative association between APGAR Score Type of Delivery, correlation coefficient -0.08 indicates that there is a weak negative association between Birth Weight and Type of Delivery and correlation coefficient 0.57 indicates a weak positive association between Mother's Age and Parity implies that relationship is not significant, which in other words implies that multicollinearity is not present in the data. It can be seen that only 2 out of 5 predictors are highly significantly to the outcome of Logistics Regression Analysis. These include Birth Weight and Type of Delivery. This means that increase in birth weight is associated with child survival at birth and increase in type of delivery is associated with child is died at birth. There is obviously a negative relationship between the status of the baby and the birth weight, i.e. weight at birth, the lower the chances that the baby will die. It also indicates that the higher the types of delivery, the higher the chances of the child dying.

The type of delivery has a direct effect on whether the child lives or not. Normal delivery has the chances of giving more life than death.

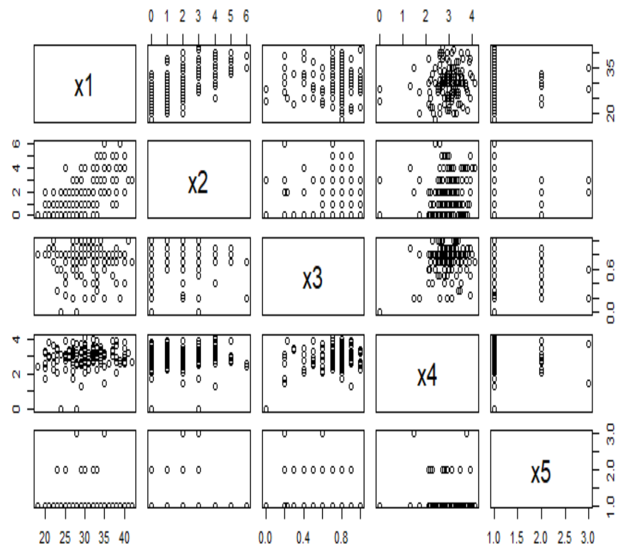

Figure: Logistic Regression of Independent Variables

4.0 Conclusion

This study uses the logistic regression model to analyse child at birth data set to investigate the factors that influence the survival of child at birth. Among the five independent variables considered, two variables (Birth weight and Type of delivery) are shown to be highly significant. Hence, the results indicate that increase in birth weight and Type of delivery are associated with child survival at birth. Therefore, this study recommend that pregnant women should be educated about the effect of baby weight in a worm as that may increases the chance of caesarean section which in turn may affect the likelihood of child survival at birth. Further studies are suggested to consider factors like educational level, income of level of the family, antennal status and blood pressure. A more advance community-based survey are recommended since not pregnant women attain formal health care facilities for antennal and postnatal services, which may expose more factors that influence child survival at birth.

Declarations

Ethics approval and consent to participate.
Not Applicable
Consent for publication
All authors have read and consented to the submission of the manuscript.
Availability of data and material
Not Applicable.
Competing interests
All authors declare no competing interests.
Funding
There was no funding for the current report.

References

Esan, D. T., Muhammad, F., Ogunkorode, A., Obialor, B., & Ramos, C. (2022). Traditional beliefs in the management and prevention of neonatal jaundice in Ado-Ekiti, Nigeria. *Enfermería Clínica (English Edition)*, *32*, S73-S76.

Emmanuel, A., Kain, V. J., & Forster, E. (2021). Improving the quality of neonatal care in Nigeria through the education of maternity health workers. *Journal of Neonatal Nursing*, *27*(3), 211-215.

Islam, M. A., & Biswas, B. (2021). Socio-economic factors associated with increased neonatal mortality: a mixed-method study of Bangladesh and 20 other developing countries based on demographic and health survey data. *Clinical Epidemiology and Global Health*, *11*, 100801.

Anandalakshmy, P. N., Talwar, P. P., Buckshee, K., & Hingorani, V. (1993). Demographic, socioeconomic and medical factors affecting maternal mortality-an Indian experience. *Journal of Family Welfare*, *39*(3), 1-4.

Adetola, A. O., Tongo, O. O., Orimadegun, A. E., & Osinusi, K. (2011). Neonatal mortality in an urban population in Ibadan, toNigeria. *Pediatrics & Neonatology*, *52*(5), 243-250.

Wilson, J. R., Lorenz, K. A., Wilson, J. R., & Lorenz, K. A. (2015). Standard binary logistic regression model. *Modeling binary correlated responses using SAS, SPSS and R*, 25-54.

Boateng, E. Y., & Abaye, D. A. (2019). A review of the logistic regression model with emphasis on medical research. *Journal of data analysis and information processing*, *7*(4), 190-207.

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia medica*, *24*(1), 12-18.

Scott, D. W. (1990). The New S Language.

Maitra, P., & Pal, S. (2008). Birth spacing, fertility selection and child survival: Analysis using a correlated hazard model. *Journal of health economics*, *27*(3), 690-705.

Audu, L. I., Otuneye, A. T., Mairami, A. B., Mukhtar-Yola, M., & Mshelia, L. J. (2021). Determination of neonatal case-specific fatality rates in a tertiary health institution in North Central Nigeria. *BMC pediatrics*, *21*(1), 1-8.

Akinyemi, J. O., Bamgboye, E. A., & Ayeni, O. (2015). Trends in neonatal mortality in Nigeria and effects of bio-demographic and maternal characteristics. *BMC pediatrics*, *15*(1), 1-12.

Adetola, A. O., Tongo, O. O., Orimadegun, A. E., & Osinusi, K. (2011). Neonatal mortality in an urban population in Ibadan, Nigeria. *Pediatrics & Neonatology*, *52*(5), 243-250.

Mesike, C. G., & Mojekwu, J. N. (2012). Environmental determinants of child mortality in Nigeria. *Journal of Sustainable Development*, *5*(1), 65.

Boerma, J. T., & Bicego, G. T. (1992). Preceding birth intervals and child survival: searching for pathways of influence. Studies in family planning, *23*(4), 243-256.

Ronsmans, C. (1996). Birth spacing and child survival in rural Senegal. *International journal of epidemiology*, *25*(5), 989-997.

Madaki, U. Y., Shu'aibu, A., & Umar, M. I. (2023). Mathematical model for the dynamics of bacterial meningitis (Meningococcal meningitis): a case study of Yobe State Specialist Hospital, Damaturu, Nigeria. *Gadau Journal of Pure and Allied Sciences, 2(2), 113–129.* https://doi.org/10.54117/gjpas.v2i2.19

Madaki, U. Y., Adamu, U. C., Muhammad, A. A., & Baba, I. A. (2023). Forecasting Infant Mortality Rate using Exponential Smoothing and Moving Averages Techniques. *Gadau Journal of Pure and Allied Sciences, 2(1), 40–49.* https://doi.org/10.54117/gjpas.v2i1.15