Impact of Socio-Demographic Factors on Survival of Under-five Children in Nigeria: Evidence from 2008 NDHS Data

Abdulmumeen A. Issa, Waheed B. Yahya and Eyitayo T. Jolayemi

Abstract—A number of discussions on mortality or survival patterns of under-five children have been presented in the literature over years, most of which were characterized by descriptive analysis, in which facts were reported by percentages, ratio and measures of association to mention a few. In this study, binary logistic regression model was employed to model the survival status (dead or alive) of under-five children in Nigeria as a function of some socio-demographic variables. Results from this study revealed that ten socio-demographic variables among several others were significantly associated with the survival of under-five children in Nigeria. Specifically, the results showed that children that were born in urban area, that were exclusively breastfed, that were among the first four children in the family, whose mothers have secondary education and post-secondary education have significant increased odds, to about 27%, 580%, 20%, 22% and 102% respectively, of surviving beyond age five than their counterparts in the reference categories of the above identified risk factors (odds ratio is 1.271(p < 0.0001 for urban), 6.810 (p<0.0001 for breastfeeding), 1.197 (p < 0.0001 for birth order), 1.225 (p = 0.001 for secondary education) and 2.023 (p < 0.0001 for higher education)). Results from this work indicated that more enlightenment program is required to stem the alarming increase in under-five mortality rate in Nigeria which currently stood at 112%. Data set from Nigerian Demographic and Health Surveys (NDHS) report for 2008 was employed in this study.

Index Terms—Mortality Rate, Under-Five Survival, Nigerian Demographic and Health Surveys, Logistic Regression Model, Odds Ratio.

I. INTRODUCTION

Under-five children mortality has been given greater attention globally and as a result, under-five mortality rate has reduced drastically, especially in the developed nations. Over the last four decades, the numbers have fallen from 20.4 million to 10.9 million annually [9]. However, the reverse is the case in sub-sahara Africa, as under-five mortality in the region moved up from 2.33 to 4.5 million annually [1], [2]. The Global steady decline during these periods notwithstanding, Nigeria still remain one of the country with highest “Infant” and “Under-five” mortality rate among the developing Nations [2]. According to Nigeria Demographic and Health survey (NDHS), 1990, 87 per 1000 infants born in Nigeria die before their first birthday while 115 per 1000 under-five children deaths was reported [4], [2]. In 1999 NDHS, an infant mortality rate of 75 deaths per 1000 live births and 140 per 1000 under-five mortality rate were estimated for the period of 1995 – 1999 [5]. In addition, a review of trends in under-five mortality rates between 1960 and 1998 by UNICEF (2000), based on an estimate of 187 deaths per 1000 for 1998 indicates a 10% reduction in Nigeria [9]. Also, in NDHS 1999, infant and under-five death rates for the ten years’ span were 70 and 102 per 1000 live birth respectively [3]. That is a reduction of about 20% and 11% respectively, from 1990 to 1999. In NPC, 2004, report, the rate was put at 100 per 1000 for infants and 201 deaths per 1000 for the general under-five mortality (as contained in NDHS data 2003) [3], [10].

Previous work on the causes of child mortality reveals that infectious and malnutrition constitute a larger percent to child mortality [11], that factors such as mother and child characteristics as well as the characteristics of the community where they live (i.e. The Socio-Demographic variables) affect the risk of infant and under-five mortality [1]. This mortality scenario has been reported to be concentrated in a poor resource setting like our nation (Nigeria) where poverty, ignorance and social instability have provide a platform on which malnutrition and infection diseases have resulted in childhood death [1].

It’s however, awful to note, that despite the efforts and huge funds voted by Nigeria government, foreign agencies and NGO’s on immunization and other interventions, to reduction the under-five mortality rate in the country [7], [8], there has been no significant improvement in the under-five survival in Nigeria. This might be due to not paying maximum attention to some socio-demographic factors like Mother’s education, Birth control, Place of resident and host of others that are hindrance to the survival of under-five children. So, as against many of the previous study on under-five children mortality whose reports are merely percentages, ratios and measures of association, our concern in the present study is to be able to quantify the effects of these factors (Socio-Demographic factors) on under-five survival by modeling of under-five survival as a function of some Socio-Demographic variables.
II. MATERIALS AND METHODS

A. Materials

1) Data used in the Study

The Data used in the write up was extracted from the 2008 edition of Nigeria Demographic and Health Survey record and recoded to suit the language of the intended method of analysis (Logistic Regression). The Response (Dependent) variable being the Survival status of the under-five children is of two categories vis. Alive or Dead. The Explanatory variables considered in the in this study are: Child’s characteristics such as, Gender, Birth order, Breast feeding, Place of delivery; Mother’s characteristics vis. Place of residence, Education and Age; wealth index; as well as Environmental factor in terms of Geo-political zones. The response variable is denoted by Y while the explanatory variables were indexed $X_j, j = 1, 2, \ldots, p$ ($p$ is the number of the explanatory variables in consideration). We define the response variable as follows:

\[ y_i = \begin{cases} 1, & \text{if the child is alive} \\ 0, & \text{if the child is dead} \end{cases} \]

2) Identification of the Variables

15 variables were used in the study; they can be identified and are coded thus:

- Status is used to denote Child’s Survival status \{1, if ‘alive’ and 0, if ‘dead’\}
- Sex is used to denote Child’s Gender \{1, if ‘Male’ and 0, if ‘Female’\}
- BirthOrd is used to denote Child’s Birth order \{1, if ‘among the 1st four children’ and 0, if ‘from 5th above’\}
- PDelivery is used to denote Child’s Place of delivery \{1, if ‘Hospital’ and 0, if ‘other places’\}
- BreastFd is used to denote Breast Feeding \{1, if ‘Exclusively Breast fed’ and 0, if ‘Otherwise’\}
- P_Reside is used to denote Mother’s place of Residence \{1, if ‘Urban’ and 0, if ‘Rural’\}
- Wealth is used to denote Wealth index of the Child’s Parent \{1, if ‘Rich’ and 0, if ‘Poor’\}
- Geo-political zone (Region) has six categories Vis.
  - North Central denoted by N_Central \{1, if ‘Living North Central’ and 0, if ‘Otherwise’\}
  - South West denoted by S_West \{1, if ‘Living in South West’ and 0, if ‘Otherwise’\}
  - North Eastern denoted by N_East \{1, if ‘Living in North East’ and 0, if ‘Otherwise’\}
  - South Eastern denoted by S_East \{1, if ‘Living in South East’ and 0, if ‘Otherwise’\}
  - South Southern denoted by S_South \{1, if ‘Living in South South’ and 0, if ‘Otherwise’\}
  - North Western
- Age@Bth is used to denote Mother’s Age at First Birth
- Primary Education is denoted by PryEdu \{1, if ‘PryEdu’ and 0, if ‘otherwise’\}
- Secondary Education is denoted by SecEdu \{1, if ‘SecEdu’ and 0, if ‘otherwise’\}
- Tertiary Education is denoted by TertEdu \{1, if ‘TertEdu’ and 0, if ‘otherwise’\}
- ‘TertEdu’ and 0, if ‘otherwise’

while, No Education is used as the reference categories

B. Method of Analysis

1) Logistic Curve/Function

Logistic regression is often used in modelling binary response variable (or general response variable), owing to the fact that: The conventional Ordinary Linear Regression (OLR) warrants the independent variable (IV) to be Quantitative in nature (taking values from measurement) while Logistic Regression is appropriate when the IV is a Discrete Choice (DC) in nature (making a choice from list of possible categories) [15]. If OLR is used to model a discrete choice variable, the predicted values may be greater than one or less the zero at some extreme values of the predictor(s) variable. Since the value of probability must lie between 0 and 1, getting such values will be a violation and will not be acceptable; The assumption that variance of the response variable ‘Y’, say, is constant across the value of the predictor variable ‘X’ (Hosomcedasticity) may not be realizable when the response variable is a DC [15]. Variance value is probability dependent; also, the normality assumption of the unobserved error term may be difficult to justify if we use OLR in modelling of Binary Variable [14]. Let $y_i$ represents the child survival status, then

\[ y_i = \begin{cases} 1, & \text{if the } i^{th} \text{ child is alive} \\ 0, & \text{if the } i^{th} \text{ child is dead} \end{cases} \]

Then, the logistics function that relates the predictor variables $X’s$ to binary variable $Y$ is given below [12], [13].

\[ P(y_i = 1 \mid x’s = 1) = \frac{e^Z}{1 + e^Z} \]

or simply:

\[ P = \frac{1}{1 + e^{-Z}} \]

\[ (1) \]

where $Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$

\[ \text{Odd}[P(y_i = 1 \mid x’s = 1)] = \frac{P(y_i = 1 \mid x’s = 1)}{1 - P(y_i = 1 \mid x’s = 1)} = e^Z \]

\[ (2) \]

Equation (1), is the conditional probability of surviving given the values of the predictor’s variable while (2), is the ratio of the prob. of survival to that of non-survival, in the present of the predictors, i.e. the Odds of survivorship.

2) Examining Model’s Goodness of fit

Although a number of goodness of fit measures are available but Likelihood Ratio Test (LRT) is the most appropriate and general test for comparing models with different level of complexity (in terms of number of Model’s parameter) [12]. LRT is asymptotically distributed chi-square, and is used in the study.

\[ LR = -2 \ln \left( \frac{L_{SM}}{L_{CM}} \right) \sim \chi^2(d_{CM} - d_{SM}) \]

\[ (3) \]

where:

- $L_{SM}$ - denotes the likelihood of the simple model (with no constant term only or fewer parameters)
\[ \kappa = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} - \sum_{i=1}^{r} n_{i} \cdot \sum_{j=1}^{c} n_{j}}{\sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij}} \sim \chi^2_{(c-1)(r-1)} \]  

(4)

where,

\[ n_i = \sum_{j=1}^{c} n_{ij}, \quad n_j = \sum_{i=1}^{r} n_{ij}, \quad N = \sum_{i=1}^{r} \sum_{j=1}^{c} n_{ij} \]

and the Tau statistic is:

\[ \tau = \sqrt{\lambda} \sim \chi^2_{(c-1)} \]  

(5)

If the calculated value in each of the two statistics is greater than the corresponding chi-square table value, then the model is sufficient and adequate enough for predicting the survival chance of under-five children.

4) Interpretation of the model’s Parameters

The contribution of each explanatory variable \(X_i\)’s will be measured in terms of percentage increase in Odds of survivorship (eqn.2) and percentage increase in the survival chance as a result of the presence of a particular predictor variable as given in (3) and (4) below respectively.

\[ \left[1 - e^{\beta_i}\right] \times 100 \]  

(6)

and

\[ \left[1 - \frac{P(y_1 = 1 | x_1 = 1)}{P(y_1 = 1 | x_1 = 0)}\right] \times 100 \]  

(7)

If the end result of the relations in 3 and 4 yield negative values, then the concerned explanatory variable increases the Odd and chance of survival. Otherwise it decreases them respectively.

III. DATA ANALYSIS

The logistic regression modeling in the study was done using of SPSS 17 and the result is summarized below. Table II give the coding of the categorical explanatory variables while Table III and IV respectively present the summarized results of the simple model (without explanatory variables) and most complex possible model (with the explanatory variables whose contributions are significant).

| TABLE I: CROSS TABULATION FORMAT OF THE PREDICTED AND OBSERVED SURVIVAL |
|-----------------------------|-----------------------------|-----------------------------|
| **Observed Status** | **Predicted Status** | **Total** |
| Dead | Alive | \(N_i\) |
| \(n_{11}\) | \(n_{12}\) | \(n_{1}\) |
| \(n_{21}\) | \(n_{22}\) | \(n_{2}\) |
| \(N_1\) | \(N_2\) | \(N\) |

Then, the kappa statistic is give as follows:

\[ \kappa = \frac{\sum_{i=1}^{c} \sum_{j=1}^{r} n_{ij} - \sum_{i=1}^{c} n_{i} \cdot \sum_{j=1}^{r} n_{j}}{\sum_{i=1}^{c} \sum_{j=1}^{r} n_{ij}} \sim \chi^2_{(c-1)(r-1)} \]

\[ \tau = \sqrt{\lambda} \sim \chi^2_{(c-1)} \]

(6) and (7)

A. Proposed model’s Parameters

The proposed model, in terms of probability is given as:

\[ P = \frac{1}{1 + e^{-\theta_0}} \]  

(8)

DOI: http://dx.doi.org/10.24018/ejers.2019.4.8.1454
for simple model, without predictors
\[ P = \frac{1}{1 + e^{-\beta_0 + \beta_1 \text{BthOrd} + \beta_2 \text{Pdel} + \beta_3 \text{Sex} + \ldots + \beta_{15} \text{P.Resd}}} \]  

(9)

for complex model, with predictors.

In terms of Odds of survival, they are as follows:

for simple model, with no predictor variables, we have:
\[ \text{Odd}(P) = \exp(\beta_0) \]  

(10)

and for complex model, with predictor variable, we have:
\[ \text{Odd}(P) = \exp(\beta_0 + \beta_1 \text{BthOrd} + \beta_2 \text{Pdel} + \beta_3 \text{Sex} + \ldots + \beta_{15} \text{P.Resd}) \]  

(11)

The estimation converges at the fifth iteration with the final -2 Log likelihood being 20060.732. Using the relation 8 and 10, we have the chance of surviving as:
\[ P = \frac{1}{1 + e^{-2.073}} \]

And the odd of surviving as:
\[ \text{Odd}(P) = \exp(2.073) = 7.949 \]

The estimation converges at the sixth iteration with the final -2 Log likelihood being 17343.627. And using the relations 9 and 11; we have the chance of surviving as:
\[ P = \frac{1}{1 + e^{-0.564 - 0.180(\text{BthOrd}) + 0.096(\text{Pdel}) - \ldots + 0.240(\text{P.Resd})}] \]

and the Odd of surviving as:
\[ \text{Odd}(P) = \exp[0.564 + 0.180(\text{BthOrd}) + 0.096(\text{Pdel}) - \ldots + 0.240(\text{P.Resd})] \]

The Log Likelihood Ratio (LLR) is 2626.105 (with P-value of 0.000) confirms the fitness of the final model in estimating the survival chance of the under-five children. The significant contribution of each explanatory variable is as given in Table V.

### Table III: Summarized Result of the Simple Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>Df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 'Constant'</td>
<td>2.073</td>
<td>.019</td>
<td>12219.814</td>
<td>1</td>
<td>.000</td>
<td>7.949</td>
</tr>
</tbody>
</table>

### Table IV: Summarized Results of the Complex and Final Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P_Resid</td>
<td>.240</td>
<td>.052</td>
<td>21.501</td>
<td>1</td>
<td>.000</td>
<td>1.271</td>
</tr>
<tr>
<td>Sex</td>
<td>-.134</td>
<td>.040</td>
<td>11.233</td>
<td>1</td>
<td>.001</td>
<td>.857</td>
</tr>
<tr>
<td>BirthOrd</td>
<td>.180</td>
<td>.041</td>
<td>18.908</td>
<td>1</td>
<td>.000</td>
<td>1.197</td>
</tr>
<tr>
<td>SecEdu</td>
<td>.203</td>
<td>.060</td>
<td>11.591</td>
<td>1</td>
<td>.001</td>
<td>1.225</td>
</tr>
<tr>
<td>TerEdu</td>
<td>.704</td>
<td>.144</td>
<td>23.861</td>
<td>1</td>
<td>.000</td>
<td>2.023</td>
</tr>
<tr>
<td>N_Central</td>
<td>.264</td>
<td>.056</td>
<td>22.224</td>
<td>1</td>
<td>.000</td>
<td>1.303</td>
</tr>
<tr>
<td>S_South</td>
<td>.169</td>
<td>.069</td>
<td>5.982</td>
<td>1</td>
<td>.014</td>
<td>1.184</td>
</tr>
<tr>
<td>Excl_Breast</td>
<td>1.918</td>
<td>.040</td>
<td>2313.211</td>
<td>1</td>
<td>.000</td>
<td>6.810</td>
</tr>
<tr>
<td>PDelivery</td>
<td>.096</td>
<td>.054</td>
<td>3.236</td>
<td>1</td>
<td>.072</td>
<td>1.101</td>
</tr>
<tr>
<td>S_West</td>
<td>.494</td>
<td>.081</td>
<td>37.208</td>
<td>1</td>
<td>.000</td>
<td>1.639</td>
</tr>
<tr>
<td>Constant</td>
<td>.565</td>
<td>.045</td>
<td>155.422</td>
<td>1</td>
<td>.000</td>
<td>1.759</td>
</tr>
</tbody>
</table>

The result of cross tabulation of the observed survival status and the predicted survival status (based on grouping individual with the predicted probability of 0.6062 and above as been ‘Alive’ and been ‘Dead’ if the predicted probability is less than 0.6062) presented in Table VI below, give Kapper and Tau statistics respectively equals 0.071 (with p-value of 0.000) and 0.1018 (with P-value of 0.000) and signifies the adequacy of the fitted model.

### Table V: Summary Contribution of Individual Variable to Survivorship of Under-Five Children

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odd ntr.</th>
<th>P1(X)</th>
<th>%CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>P_Resid(Urban)</td>
<td>0.240</td>
<td>1.271</td>
<td>27</td>
</tr>
<tr>
<td>Sex(Male)</td>
<td>-.134</td>
<td>.857</td>
<td></td>
</tr>
<tr>
<td>PDelivery(Hosp)</td>
<td>.096</td>
<td>1.197</td>
<td>20</td>
</tr>
<tr>
<td>SecEdu(1)</td>
<td>.203</td>
<td>1.225</td>
<td>23</td>
</tr>
<tr>
<td>TerrEdu(1)</td>
<td>.704</td>
<td>2.023</td>
<td>102</td>
</tr>
<tr>
<td>Excl. Breast</td>
<td>1.918</td>
<td>6.810</td>
<td>581</td>
</tr>
<tr>
<td>P_Del(1)*-4b</td>
<td>.180</td>
<td>1.197</td>
<td>20</td>
</tr>
<tr>
<td>N_Central(1)</td>
<td>.264</td>
<td>1.303</td>
<td>30</td>
</tr>
<tr>
<td>S_South(1)</td>
<td>.169</td>
<td>1.184</td>
<td>18</td>
</tr>
<tr>
<td>S_West</td>
<td>.494</td>
<td>1.639</td>
<td>64</td>
</tr>
<tr>
<td>Constant</td>
<td>.565</td>
<td>1.759</td>
<td></td>
</tr>
</tbody>
</table>

P(1) is the probability of surviving given that the value of the predictor variables is zero
P(1/X = 1) is the probability, of surviving given the value of the predictor variable of interest is one
%CP(1) is the percentage increase in the survival chance when the value of the predictor variable of interest is one.

**IV. DISCUSSION OF RESULTS AND CONCLUSION**

- 10 out of the 15 socio-demographic variables considered were retained in the final fitted model been the significant contributors, as shown in table IV. Meanwhile the remaining 5 vis.: “N_Eastern Region; S_Eastern Region; Primary Education; Wealth Index; and Mother’s Age at first Birth” does not contribute significantly to the Model when other variable are already included. However, “Tertiary level of Mother’s Education”, “Exclusive Breast Feeding” and “Living in the South Western Region” are the very strong and important factor that influence Under-five children survival. They increase the odds of survival to about 102%, 581% and 64% respectively of other children in the reference categories (Table V, column 4)

- The Under-Five Mortality rate, according to the available NDHS 2008 data is given as 112 deaths per thousand life birth, which is an appreciable gain, when compared with 115, 140 and 201 as reported in NDHS, 1990 [4] and 1999 [5] and NPC, 2004 [6] respectively. In order words, survival rate moves up.
REFERENCES


[2] UNICEF. Infant and under-five mortality (s)


[8] Executive board room; sixth meeting of the technical group on the global eradication of poliomyelitis. WHO Geneva draft meeting, 2001:1-6 (s)


[10] National Demographic and Health Survey, 2003 (s)


