IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 11, Issue 4 Ver. I (Jul - Aug. 2015), PP 47-52
www.iosrjournals.org

Modelling the Determinants of Maternal Mortality: A Comparative Study of Logistic Regression and Artificial Neural Network Model

Chukwu A.U [1*] & Oladeji R.T [2]
[1.] *Department of Statistics, University of Ibadan, Ibadan, Nigeria*
[2.] *Dept of Mathematics, Federal college of Education, Katsina, Nigeria*

Abstract: *Without any doubt, Indicators derived from mortality rates give a clear representation of the overall population health. One of such indicators is maternal mortality. This work discusses Logistic regression and Artificial Neural Network model and the application of these models in predicting maternal mortality. 276 records (ranging from 2003 to 2012) on mother's age, mode of delivery, parity, sex of the baby, baby's weight at birth, nature of complication (independent variables) and mother's status (dependent variable) were collected from the medical record department of the University College Hospital Ibadan. Logistic regression model was used to check for the risk factor associated with maternal mortality. In order to compare the efficiency of ANN and logistic regression model, the following measures were used: sensitivity, specificity and goodness of fit. Results of the analysis revealed that parity and age are the major determinant of maternal mortality. The result of the comparison of the efficiency of ANN and logistic regression model showed that ANN outperformed logistic regression with sensitivity 50.6% versus 31.0%, specificity 91.6% versus 86.6% and the mean square error (MSE) of ANN is very small compared to that of the logistic regression model.*
***Keywords**: Maternal mortality, Artificial Neural Network, Logistic Regression.*

I. Introduction

Nicholas (2007) describes mortality as that branch of Demography which deals with the total process of deaths and the changes it brings about in the population. Fabayo (2010) defines maternal mortality as death of a woman while pregnant or within 42 days of termination of pregnancy or delivery from any cause related to or aggravated by the pregnancy or its management. The World Health Organization (WHO 2005) in its article "make every mother and child count" defines maternal mortality as the death of a woman while pregnant or during delivery or within 42 days of termination of pregnancy irrespective of the duration and site of pregnancy, from any cause related to or aggravated by the pregnancy or its management. All maternal deaths are subdivided into two, direct and indirect maternal death (Nicholas 2007). Direct maternal death is as a result of complication during pregnancy or delivery or management of both, while indirect maternal death is a pregnancy related death in patient with a preexisting or newly developed health problem unrelated to pregnancy. This indirect maternal death is referred to as accidental, incidental or non-obstetrical maternal death.

Maternal mortality is caused by a number of factors among these are age of the mother, parity (number of children a woman has), weight of the baby at birth, method of delivery (this could be through SVD, CS or FORCEPS), birth status, complications like eclampsia, sepsis, obstructed or prolonged labor, post partum and anti partum hemorrhage etc. Every minute, a woman dies during labor or delivery. The highest maternal mortality rates are in Africa with a lifetime risk of 1 in 16, the lowest rates are in Western nation (1 in 2,800) with a global ratio of 400 maternal deaths per 100,000 live births (Nawal 2008). Akinola, a member of society of gynecologist and obstetrician cited by Olanrewaju (2013) disclosed that 145 Nigerian pregnant women die daily either during childbirth or shortly after. The tragedy is that these women die not from diseases but during normal life enhancing process of procreation. Most of this death can be avoided if preventive measures were taken and adequate care was available. Safe motherhood human rights issue states that 'the death of a woman during pregnancy or childbirth is not only on health issue but also a matter of social injustice' (Fabayo 2010).

The high maternal mortality rate in the country has been a cause of concern for many people, especially with respect to meeting the fifth millennium development goal 2015 which is reduction of maternal mortality rate by three quarter (75%). According to Akinola, cited by Olanrewaju (2013), he stated that Nigeria high maternal mortality rate is second only to that of China adding that the country held the record as Africa's highest contributor to MMR, also the CIA World Fact Book (2013) puts MMR in the country at 630 per 100,000 live births. Government at both state and federal levels allocate huge sum of money to health care and still MMR is high in the country, in terms of health infrastructure Nigeria is having about 18,258 PHC facilities, 3,275 secondary facilities and 29 tertiary facilities (NHMIS 2002). With the efforts so far made by the government to bring MMR down, it is still not yielding the desired result; it is therefore of reasonable interest to identify some underlying factors that may be responsible for high maternal mortality rate, study the extent to which the factors may influence maternal mortality and look for an efficient statistical model that can predict maternal mortality.

1.1 Logistic Regression and Artificial Neural Network model

Logistic Regression (LR) and Artificial Neural Network (ANN) models are non linear regression models used for analyzing and modeling. These models are intrinsically non- linear and can only be linearised by a numerical search procedure. Logistic Regression is a type of regression analysis that is used for predicting the outcome of a categorical dependent variable based on one or more predictor variables. The probabilities describing the possible outcome of a single trial are modeled as a function of explanatory variables using a logistic function. Logistic regression can be binomial or multinomial, binomial or binary logistic regression refers to the instance in which he observed outcome can have only two possible types (e.g. dead or alive, yes or no, etc.). Multinomial logistic regression refers to cases where outcome can have three or more possible types (e.g. better, no change, worse.). In binary logistic regression, the outcome is coded 0 and 1 as it leads to the most straight forward interpretation. Like other regression analysis, the logistic regression makes use of one or more predictor variables that may be either continuous or categorical data and also the expected value of the response variable is fit to the predictors. The general form of a logistic regression model is given as

$$Yi = \frac{\exp(\beta_0 + \beta_i \sum_{i=1}^{n} x_i)}{1+\exp(\beta_0 + \beta_i \sum_{i=1}^{n} x_i)} + \ell_i$$

Where β_0 is the intercept

B_i is the parameter

x_i are the explanatory variables

ℓ_i is the random error and Yi the response variable.

Everitt and Skrondal (2010) defined Artificial Neural Network as a mathematical structure modeled on the human neural network and designed to attack many statistical problems, particularly in the areas of pattern recognition, multivariate analysis, learning and memory. The essential feature of such a structure is a network of simple processing elements (artificial neurons) coupled together so that they can cooperate. From a set of 'inputs' and an associated set of 'parameters', the artificial neurons produce an 'output' that provides a possible solution to the problem under investigation. ANN has three layers, input, output and hidden layers. In many ANN model the relationship between the input received by a neuron and its output is determined by a generalized linear model.

ANN is typically defined by three parameters:

(a) The interconnection pattern between different layers of neurons.
(b) The learning process for updating the weights of the interconnections.
(c) The activation function that converts a neuron's weighted input to its output activation.

The only difference between LR and ANN is the presence of hidden neurons present in the hidden layers. This makes ANN perform better than LR model. So many authors have worked on the comparison of the predictive ability of logistic regression and ANN models, the following are the review of related literatures on the comparison of the predictive ability and accuracy of logistic regression and ANN models. Sharareh R. (2014) used LR and ANN to predict type of pregnancy in women equal or greater than 35 years of age, data of 1404 women of age 35 or more were collected and the result of the classification task showed about 82% accuracy, 76% specificity, 56% sensitivity and the area under the ROC 0.67(67%). This result shows that ANN is a better predictive model. Eftekhar. B et al (2005) used ANN and Logistic Regression models to predict mortality in head trauma based on initial clinical data. It was found out that ANN significantly outperformed logistic model in both fields of discrimination and calibration but under performed in accuracy. Shi H.Y et al (2012), compared ANN and logistic regression model for predicting In Hospital mortality after primary cancer surgery, this study retrospectively compared 1,000 pairs of logistic regression and ANN models, based on the initial clinical data for 22,926 HCC surgery patients compared to the logistic regression model, the ANN had a better accuracy rate in 97.28% of cases, a better HCC statistics in 41.18% and a better AUROC curve in 84.67% of cases. This showed that the ANN was more accurate in predicting mortality and had higher overall performance.Lin CC et al (2010) also compared the performance of ANN and LR for predicting mortality in elderly patients with hip fracture and found out that ANN outperformed LR.Valerie B et all (2010) also used ANN and logistic regression models as classification models for variable selection for prediction of breast cancer patients outcome and it was found out that in sensitivity, specificity and area under the ROC, ANN outperformed logistic regression model.

II. Data And Methods

The data used for this research work were extracted from the medical record department of the University College Hospital (UCH), Ibadan. Two hundred and seventy six records (ranging from 2003 to 2012) were collected and these were subdivided into two, the in-sample and out-sample data set. The in sample data set were used to build the model and this is made up of the first 200 records while the out sample data set were used to test or validate the model and this is made up of the last 76 records. The records collected were on the following items:

a. Mother' age.
b. Parity: Number of children that a woman has.
c. Method of delivery: This could be through spontaneous virginal delivery (SVD), surgical operation (CS), or through forceps delivery.
d. Sex of the baby.
e. Baby's weight at birth.
f. Nature of complication encountered by the mother: Examples are eclampsia, obstructed labor or prolonged labor, sepsis, post partum and anti partum hemorrhage, breech or transverse presentation, anemia, etc.
g. Mother's status at birth or shortly after birth.

Items a to f serves as the independentvariables while item g is the dependent variable.

2.1 Metrics for Comparison of Logistic Regression and Artificial Neural Network Model

Logistic regression model was used to check for the risk factor(s) associated with maternal mortality. Feed forward ANN with logistic transfer function and Logistic regression models were compared using same training and testing data sets, the ANN used is made up of eight parameters and three hidden layers. While logistic regression has six parameters. The result of the two models were compared using the following metrics:

a. Sensitivity: This is the proportion of the cases having positive test result (true positive) of all positive cases tested.
b. Specificity: This is the proportion of true negative of all the negative cases tested
c. PredictiveValues: This is a measure of the times that the value positive or negative is the true value. Positive predictive indicator consists of the percentage of people with a positive diagnostic test result and negative predictive indicator consists of the percentage of people with a negative diagnostic test result.

$$PPVi = \frac{TPi}{TPi + FPi} \quad \text{where } TPi + FPi \neq 0$$
$$NPVi = \frac{TNi}{TNi + FNi} \quad \text{where } TNi + FNi \neq 0$$

Confusion matrix table

Actual Value	Predicted Value	
	Y	N
Y	TP	FN
N	FP	TN

d. Goodness of Fit: This provides an overall measure of the fit of the model and is usually not sensitive when the fit is poor for just a few models.

2.2 Methodology: To check for the risk factor associated with maternal mortality using logistic regression analysis, stepwise regression analysis was used where we have a full model (with all the independent variables and the intercept), model with all the independent variables and no intercept and the last model is the reduced model (only the significant variable), this is given in table 1 below.

Odd ratios were also estimated for only the significant variables as given in table 2.

Table 1: Logistic Regression Model with Significant Predictor Variable.

Coefficient	Estimate	Std Error	Z Value	Pr(> |Z|)	Profile Confidence Interval (Coefficient)	
					2.5%	97.5%
Age	0.08222	0.02704	3.040	0.00236 **	0.03088	0.13737
Parity 0	-3.01108	0.78560	-3.833	0.00013***	-4.62540	-1.53287
Parity 1	-2.45275	0.08707	-2.817	0.00484**	-4.21850	-0.79104
Parity 2	-2.32171	0.93753	-2.476	0.01327*	-4.22909	-0.53995
Parity 3	-3.14788	1.02163	-3.081	0.00206**	-5.22902	-1.20574
Parity 4	-3.50009	1.11506	-3.139	0.00170**	-5.76604	-1.37597
Parity 5	-3.93931	1.63565	-2.408	0.01602*	-7.57441	-0.84159
Parity 6	-19.10132	1455.39800	-0.013	0.98953	-	276.10125
Parity 8	-19.18353	1455.39802	-0.013	0.98948	-	275.95506

Null Deviance: 277.26 on 200 degrees of freedom
Residual Deviance: 250.03 on 191 degrees of freedom

Log-likelihood:-125.015
AIC: 268.03

From table 1 we see that parity 0, age, parity 3 and 4 are significant that is, these factors contribute more to maternal mortality when compared to other factors considered in this research work.

Table 2: Estimated odd ratios for factors age and parity (only significant level of parity were considered)

Coefficient	Odd Ratio	Confidence Interval (95%)
Age	1.08569	1.0314, 1.1473
Parity 0	0.04924	0.0098, 0.2159
Parity 1	0.08606	0.0147, 0.4533
Parity 2	0.09811	0.0146, 0.5828
Parity 3	0.04294	0.0054, 0.2995
Parity 4	0.03020	0.0031, 0.2526
Parity 5	0.01946	0.005, 0.4310

The odds of survival of a mother during or after birth due to the factor parity increases slightly as it's level increase from 0 to 2, and then decrease slightly from levels 3 to 5 (only significant levels of parity were considered). The profile confidence intervals for these significant levels of parity are [0.00098, 0.2159], [0.0147, 0.4533], [0.0146, 0.5828], [0.0054, 0.2995], [0.0031, 0.2526] and [0.0005, 0.4310]. Also for factor age the odd of a mother surviving is 1.08569 with confidence interval [1.0314, 1.1473].

Table 3: Goodness- of-fit measures for both the in and out sample data analysis.

In sample statistic			Out sample statistic	
ME	MAE	RMSE	MAE	RMSE
-0.02224	1.98368	2.12981	0.49752	2.09558

The estimated Logistic regression model for mother's status is given below

$$\frac{\exp(0.082\,\mathrm{Age}i - 3.011\,P0i - 2.453\,P1i - 2.322\,P2i - 3.148\,P3i - 3.500\,P4i - 3.939\,P5i - 19.101\,P6i - 19.184\,P8i)}{1+\exp(0.082\,\mathrm{Age}i - 3.011\,P0i - 2.453\,P1i - 2.322\,P2i - 3.148\,P3i - 3.500\,P4i - 3.939\,P5i - 19.101\,P6i - 19.184\,P8i)}$$

Table 4: In and Out sample prediction/tracking.

	In sample evaluation			Out sample evaluation		
Actual mother's status	Fitted mother's status			Fitted mother's status		
		0	1		0	1
	0	103 (TN)	16 (FP)	0	26 (TN)	3 (FP)
	1	56 (FN)	25 (TP)	1	40 (FN)	7 (TP)

From table 4, the fitted model was able to track correctly 103 cases of mother's status 0 and 25 cases of mother's status 1. The incorrect predictions are 56 cases of status 0 and 16 cases of status 1. Also the fitted model was able to track correctly 26 cases of status 0 and 7 cases of status 1; the incorrect tracking is 3 cases for status 0 and 40 cases for status 1 for out sample prediction.

The sensitivity and specificity of LR model are 31.0% and 86.6% for the in-sample statistic and 14.9% and 89.7% for the out-sample statistic.

ANN Model: ANN with logistic transfer function was constructed. The network consisted of an input layer, a hidden layer and an output layer. Three ANN network were constructed, 8-2-1, 8-3-1, 8-4-1 and 8-5-1, with 21, 31, 41 and 51 weights respectively. The model has RMSE 0.44109, 0.42547, 0.41513 and 0.40399. The last model (8-5-1) was selected to fit mother's status for both the in sample and out sample data set and the prediction is shown in table 6 below.

Table 5: Goodness-of-fit measure for ANN

	In sample			Out sample	
	ME	MAE	RMSE	MAE	RMSE
ANN: 8-5-1	-0.00012	0.33115	0.40399	0.51792	1.86095

Table 6: In sample and Out sample tracking/prediction

Actual mother's status	In sample evaluation			Out sample evaluation		
		Fitted mother's status			Fitted mother's status	
		0	1		0	1
	0	109 (TN)	10 (FP)	0	24 (TN)	5 (FP)
	1	40 (FN)	41 (TP)	1	39 (FN)	8 (TP)

From table 6 the fitted model is able to track correctly 109 cases of mother's status 0 and 41 cases of status 1. The incorrect tracking is 40 cases of status 0 and 10 cases of status 1for the out sample the model is able to track correctly 24 cases of status 0 and 8 cases of status 1. The incorrect tracking is 39 cases of status 0 and 5 cases of status 1.

The sensitivity and specificity of ANN are 50.6% and 91.6% for the in sample statistic and 17.0% and 82.8% for the out sample statistic.

III. Discussion Of Result

From table 1 of the logistic regression analysis, parity; 0, 4, 3 and age are more significant with p values 0.000127, 0.001696, 0.002062 and 0.002362. This is to show that these factors contribute more to maternal mortality.

3.1 Summary of Goodness-of-fit Measure of LR and ANN

From table 3 and 5 (in sample evaluation), the MSE of LR and ANN models are 2.12981 and 0.40399 respectively and also from the out sample evaluation the MSE of LR and ANN models are 2.095578 and 1.860952. From this we see that the goodness mean square error (MSE) of ANN model is lesser than that of Logistic regression model.

3.2 In sample and out sample tracking performance of LR and ANN (table 4&6)

The actual mother's status for status 0 is 119 out of 200 and for status 1 is 81 out of 200 for the in sample data. While for the out sample data the actual status for 0 is 29 out of 76 and for status 1 is 47 out of 76.

The proportions of correct tracking/predictions by the Logistic regression and ANN models are 103/200 and 109/200 respectively for mother's status 0.The proportions of correct tracking/predictions for mother's status 1 are 25/200 and 41/200 for Logistic regression and ANN models respectively.

Also the proportions of incorrect tracking/predictions for mother's status 0 is 16/200 and 10/200 for LR and ANN models. While for status 1 we have 56/200 and 40/200 for LR and ANN modelthe proportions for correct tracking/predictions for status 0 are 26/76 and 24/76 for LR and ANN models and for mother's status 1 are 7 /76 and 8/76 for LR and ANN models respectively. While the proportions of incorrect tracking/predictions for mother's status 0 are 3/76 and 5/7 for LR and ANN and for mother's status 1 are 40/76 and 39/76 for LR and ANN respectively. Here we see that for the in sample evaluation ANN out performed LR model but in the out sample where we have smaller data set there is no significant difference in their performance, this shows that in a larger data set ANN is more efficient

3.3 Sensitivity and Specificity of LR and ANN:

The LR and ANN has sensitivity 31.0% versus 50.6%, specificity 86.6% versus 91.6% for the in sample evaluation and sensitivity 14.9% versus 17.0%, specificity 89.7% versus 82.8% for the out sample evaluation. The sensitivity and specificity of ANN is better than that of LR in the in sample evaluation, for the out sample evaluation ANN has a better sensitivity and LR has a better specificity.

IV. Conclusion

The result of the Logistic regression analysis revealed that parity (the number of children that a woman has given birth to) and age are the major determinant of maternal mortality, it was discovered that parity 0 (first pregnancy) are at a higher risk of having complication that may lead to death. This reduces as the number of children increases to 2 and increases as the number increases to 3 and 4, thus we conclude that among all the factors considered for this analysis parity and age are more significant. From the comparison of the efficiency of LR and ANN model, the LR and ANN have sensitivity 31.0% and 50.6%, specificity 86.6% and 91.6% respectively. In both the in sample and out sample data analysis we see that the mean square error (MSE) for the ANN model are very small compared to those of the Logistic regression model. Thus, we conclude that the ANN model fits the data more adequately than the Logistic regression model and hence produced better forecast results

V. Recommendation

From the result of the analysis of Logistic regression and ANN models, it is recommended that:

a. Women during their first pregnancy (Parity 0) should be given special care and attention during the antenatal visit as they at a higher risk of having complications that may lead to maternal mortality.
b. Women should be advised to give birth to their children within the reproductive age 15 to 49 as recommended by the World Health Organization (2005) as this will help in reducing maternal mortality.
c. Since there is a need to devise more robust and efficient prognostic and predictive models for maternal mortality due to the increase choice of better quality of life and also to achieve the fifth goal of the MDG

(2015), which is reduction of maternal mortality by three quarter (75%), this research recommend the construction of ANN model for prediction of maternal mortality as this model is feasible, adaptable and useful to gynecologist, mothers and the general public.

References

[1]. Central Intelligence Agency (CIA): retrieved from http://www.cia.gov/library/publications
[2]. Eftekhar B. et al (2005) 'Comparison of ANN and Logistic Regression Model for Prediction of Mortality in Head Trauma Based on Clinical data. Available at 'http://www.biomedcentral.com/1472_6947/15/3.
[3]. Everitt B.S and Skrondal A (2010) 'The Cambridge Dictionary of Statistics'
[4]. 5th Edition. Cambridge University Press, New York. ISBN-13 978-0-511-78827-7. Available at www.cambridge.org/9780521766999.
[5]. Fabayo C.O (2010) 'Maternal and Child Health for Health Workers.' Samadex Printing Press, Ilorin, Nigeria. ISBN 978_187_301_9.
[6]. Federal Republic of Nigeria: 'National Health Management Insurance Scheme'(2002).
[7]. Lin C. C. et al (2010). 'Comparison of ANN and Logistic Regression Model for Predicting Mortality in Elderly Patients with Hip Fracture.' Dept of orthopedics. National Taiwan University Hospital. Taiwan.
[8]. Nawal M. N. (2008) 'An Introduction to Maternal Mortality. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2505173.
[9]. Nicholas N. (2007).' A handbook of Methods and Measures in Demography.' ACADEC Press Legon . Ghana.
[10]. Olanrewaju S. (2013). 'Trends in Maternal Mortality'. Tribune.com.ng/ en/…/1117/_ mortality_and_race_to_2015.
[11]. Sharareh R. (2014). Artificial Neural Networks Application to Predict type of pregnancy in women equal or greater than 35 years of age. International Journal of computer and information technology (ISSN 2279-0764). Vol 3-issue 1.
[12]. Shi H. Y. (2012)'Comparison of Logistic Regression and ANN for Predicting In Hospital Mortality after Primary Cancer Surgery'., Nov, 16(11):
[13]. 2126_31.doc10.1007/s11605_012_1986_3.
[14]. The World Health Report (2005): Make Every Mother and Child Count. Geneva Switzerland: WHO.(2005). Retrieved from http://www.who.int /whr/2005_en.pdt.
[15]. Valerie B. et al (2010). 'Comparison of ANN with Logistic Regression Model as Classification Models for Variable Selection for Prediction of Breast cancer Patients Outcomes'. Advances in ANN System. Vol 2010(2010), ID309841.
[16]. WHO (2013), MDG5 www.who_int/topics/maternal_health/en/.