Regression and Neural Networks Analysis in Vesco-Vaginal Fistula Causality: A Comparative Approach

[1]James, T.O, [2]Udomboso, C. G and [1]ONWUKA G. I
[1]Department of Mathematics, Kebbi State University of Science and Technology
[2]Department of Statistics, [2]University of Ibadan, Ibadan, Nigeria

Abstract
Vesico vaginal fistula (VVF) is an abnormal opening of the vaginal wall to the bladder or rectum resulting in the leakage of urine. It is one of the worst morbidities associate with delivery and is a major public health problem on the rise with an estimated minimum of 150,000-200,000 patients in Nigeria. Neural network are able to solve the nonlinear regression problem. Very little research has been conducted to model the causes of VVF using artificial neural networks. The data set obtained from the case records of women admitted with cases of Vesico-vaginal Fistula (VVF) in Maryam Abacha Women and Children Hospital Sokoto, from January 2000 to December 2010 was used. We then compared the performance of Statistical neural networks and Regression model. In comparison to traditional methods, the value of Obstructed labour and misuse of instrument in ANN has higher R square (0.8 & 0.54) in which is a better result, lower MSE (2011 &4579.6) which is also a better result. The p-value is only greater than 0.05 in obstructed labour. The results of the t and F statistics confirms the better performance, since any p-value lesser than 0.05 shows that that cause of VVF cases is very significant. Therefore, we can accept the fact that MISUSE OF INSTRUMENT and YANKAN GISHIRI are both significant to cases of VVF, using ANN, while LR is not since the R squares are low. Statistical neural network model showed better predictions than various regression models for causes of VVF. However, both methods can be used for the prediction of causes of VVF.
Keywords: *Vesicovaginal Fistula, Linear Regression (LR), Artificial Neural Networks (ANN),*

1.0 Introduction

Vesico-vaginal Fistula (VVF) occurs when the blood supply to the tissues of the vagina and the bladder is restricted during prolonged obstructed labour, the tissues die between these organs, forming holes through which urine can pass uncontrollably. While Recto-vaginal Fistula (RVF) occurs in a similar way to VVF however, holes form between the times of the vagina and rectum, leading to uncontrollable leakage of feaces (**Rodney et-al, 1997**). In the case of obstetric fistula it is result of pressure exerted by foetal head in the pelvis during obstructed labour, a force that interrupted the blood flow to nearby tissues in the mother's pelvis, resulting in two classifications vesico vaginal and recto vaginal fistula. Obstetric fistula remains a challenge to the surgeon and a difficult socially unacceptable nuisances to patients, more so if it recurrent Chilton (1998). In developed countries the most common cause of VVF remains iatrogenic injury during Gynecological surgery (Godwdin and Scordino (1980). Obstetric VVF due to obstructed labour has long been eradicated from the developed world (Aarrowsmith, 1991). While in developing countries fistula and other urogenital fistula are estimated to occur in 2 percent of obstructed labours and labour can take days to resolve. Prolong compression of maternal soft tissue between fetal head and maternal pelvis leading, resulting in urogenital, fistula (Alan 2008), but it remains a major problem in developing countries like Nigeria.

Obstetric fistula is a major public health problem on the rise with an estimated minimum of 150,000-200,000 patients in Nigeria. The statistics in Northern Nigeria is even higher than the national average as a result restrictions leading to poor accessibility to health services. A total of 20,000 fresh cases of VVF were recorded annually in Nigeria and that 90 percent of the cases were untreated. The high maternal mortality and morbidity have been associated with these factors along with harmful traditional practices and other socio-cultural related to pregnancy and child birth in Northern Nigeria. (Sambo, 1990). Obstetric fistula is one of the most important reproductive health problems that need special attention especially in Northern Nigeria. It affects numerous girls and women every day and the condition leaves these affected women in a state of despair. Women so affected have to suffer not only the consequences of loosing their children but also are subject to social humiliation, shame and embarrassment. They may become outcasts due to pungent smell and wetness from urinary or feacal incontinence (Galati 2011).

The World Health Organisation (WHO) had described fistula as "the single most dramatic aftermath of neglected childbirth". The word "fistula" is a collective medical term for any abnormal connection between two bodily organs. In the case of obstetric fistula it is the result of pressure exerted by the foetal head in the pelvis during obstructed labour, a force that interrupts the blood flow to nearby tissues in the mother's pelvis, resulting in two classifications Vesico-vaginal Fistula and Recto vaginal Fistula. In this paper Vesico-vaginal Fistula was study. Vesico-vaginal Fistula represents a significant morbidity in female urology Raza et.al (2005).

A neural network is an information processing system that is inspired by the way the biological nervous system operates. Hence a neural network process information in similar manner to how the brain would process information. Neural Networks can be thought of as a machine that is designed to stimulate a particular way the human brain performs a particular task (Haykin 1995).The basic processing elements of neural networks are called *artificial neurons*, or *simply neurons* or *nodes*. In a simplified mathematical model of the neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons is represented by a transfer function. In recent times, Artificial Neural Networks (ANN) has become a widely used method. Artificial intelligence has been used in medicine for the clinical functions of diagnosis, prognosis and survival analysis, and decision support. It has been used in a wide variety of medical domains such as oncology, critical care, tuberculosis, cardiovascular and renal transplantation. A common task in medicine is thus classification using predictive models. (Taryn⊔2006). Artificial neural networks have been researched and used for applications in many different fields. Many of these areas are using ANNs to solve problems previously thought to be impossible or very difficult with traditional methods. These include face recognition, prediction of time series events, function approximation, process optimization, and others (Cheng and Titterington, 1994). There are many reasons for the amount

and robust. Unlike regression, where a specific equation must be predetermined based on the data in the system to relate the input to output variables, the general structure of an ANN can be applied to practically any system (Zealand, 1999). Also, ANNs have been shown to outperform regression models when outliers exist in the data (Denton, 1995).

Sargent (2001) performed a literature review on approximately thirty articles comparing ANNs to statistical regression for biomedical applications and found that ANNs outperformed statistics in only ten of the cases. The other articles either found that both models had equivalent performance or that regression models were better. In applications with large sample sizes, it was found that ANNs never performed better than regression. He speculated that the reason ANN did not dominate over statistical models is because both are heavily limited by the data being collected, in terms of the amount of data and the amount of error or noise in the data. In this study, the comparison of neural network model and linear regression model on the of causes of VVF in relation to Aetiological factor (prolong obstructed labour, Yankan Gishiri or Gishiricut and misuse of instrument) to determine the accurate values of multiple regression models and neural network model

1.1 Causes Of Vesicovaginal Fistula

Direct causes of vesico vaginal fistula (VVF) include the prolonged unrelieved obstructed labour, accidental Surgical injury related to pregnancy; crude attempt induced criminal abortion; female genetic mutilation (FGM). The female circumcision or female mutilation is a cultural practice that profoundly affects the health of the women it is estimated that in Africa, over 8 million women are circumcised. Also, the Yankan Gishiri or Gishiricut: This is a traditional surgical cut performed on women diagnosed to be suffering from Gishiri disease. A traditional leader or traditional birth attendants (TBA) perform the operation. The word Gishiri is a general term applied by the Hausa's to a wide range of ailments related to the productive system. The TBA's use the 'gishiri cut' to widen the birth carnal of the women to increase the passage for the baby, using a razor blade. This is also utilized as a means of stopping bleeding that could occur at childbirth. Where this is undertaken lightly it results into damage to the birth canal but it may also damages the bladder and the rectum resulting into VVF. While the indirect causes are poverty, particular of girls, early marriage: many girls become pregnant in their early teens before their pelvis is fully developed. They have high risk of obstructed labour and ultimately VVF or maternal death.

2.0 Materials And Method

2.1. Study Area

This study was carried out on data on VVF patients admitted in Maryam Abacha Women and Children Hospital Sokoto, one of the seven VVF- centers that was established in Northern Nigeria where at least 150-200 repairs are being performed each year. From January 2000 to December 2010. The case records

2.0 Materials And Method

2.1. Study Area

This study was carried out on data on VVF patients admitted in Maryam Abacha Women and Children Hospital Sokoto, one of the seven VVF- centers that was established in Northern Nigeria where at least 150-200 repairs are being performed each year. From January 2000 to December 2010. The case records of women admitted with cases of Vesico-vaginal Fistula (VVF). The following information were obtained from patient's medical records Geographical location, Duration of leakage, Aetiological factors of VVF, Age of patients, outcome of delivery, parity of patients at presentation, marital status of the patients, place of delivery, complication at presentation, Educational status of the patients, psychosocial effects.

2.2 Linear Regression

Statistical regression is a method that has been used by statisticians and engineers for many years to fit an equation to a set of data. That is Regression analysis is a technique for modeling the relationship between two or more variables. Regression models quantitatively describe the variability among the observations by partitioning an observation into two parts (Alan and Man 1982 and Montgomery 2001).

We recall the linear regression model (*LRM*) given as:

$$y_i = f(x_j, \beta) + e_i; \qquad i = 1, 2, \ldots, n \tag{1}$$

which is made up of the predicted part and the residual part. The residual is the difference between the observed and the predicted values which is ascribed to unknown sources. n is the number of observations, y_i is the i^{th} observation, $x_j = (x_{1i}, x_{2i}, \ldots, x_{ki})$ is the predictor variable vector related to y_i, $\beta = (\beta_0, \beta_1, \ldots, \beta_p)$ is the parameter vector, and e_i is the error associated with ith observation.

Explicitly, this is written as

$$y_i = \alpha + \beta x_i + e_i; \qquad i = 1, 2, \ldots, n \tag{3}$$

where y_i is the dependent variable, x_i is the independent variable (in this case, the '$years$'), α is the intercept, β is the parameter associated with the independent variable, x_i, and e_i is the stochastic term or error associated with the model.

We minimize (3) with respect to α and β, $\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \alpha}$ and $\frac{\partial \sum_{i=1}^{n} e_i^2}{\partial \beta}$, to obtain two normal equations respectively. Solving the normal equations, we obtain the estimates of the parameters α and β:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \tag{4}$$

$$\hat{\beta} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{5}$$

The predicted model becomes

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i \tag{6}$$

and the residual is given as

$$e_i = y_i - \hat{y}_i \tag{7}$$

2.3 Statistical Neural Network

In general, neural networks are able to solve the nonlinear regression problem. The statistical neural network (SNN) model structurally is composed of two parts: the predictive and the residual, as is in classical regression, given as

$$y = f(X, w) + e_i \tag{8}$$

Where $f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h g(\sum_{i=0}^{I} \gamma_{hi} x_i)$. Thus equation (6) can be written as

$$y = \alpha X + \sum_{h=1}^{H} \beta_h g(\sum_{i=0}^{I} \gamma_{hi} x_i) + e_i \tag{9}$$

$X = (x_0, x_1, \ldots, x_I)$ is the vector of the input variable, $g(.)$ is the transfer (or activation) function and $w = (\alpha, \beta, \gamma)$ are the weights (or parameters) associated with the input vector, hidden neuron and the transfer function respectively, while e_i is the error associated with the network. We note that when there is no hidden neuron, the SNN reduces to the ordinary regression model.

The weights are estimated using Taylor's first order approximation,

$$y_i = y_i^0 + \frac{\partial f(x_i, w)}{\partial w}\Big|_{w=w^0}^{(w-w^0)} + e_i \tag{11}$$

where $y_i^0 = f(x_i, w^0)$

if $\theta = w - w^0$, and $z = \frac{\partial f(x_i, w)}{\partial w}$, then we can write equation (6) as

$$y_i^* = z_i\theta + e_i \tag{12}$$

where $y_i^* = y_i - y_i^0$

The least squares estimate of the parameter θ is

$$\hat{\theta}' = (Z'Z)^{-1}Z'Y \tag{13}$$

and the estimated model is

$$\hat{Y}^* = X\hat{\theta} \tag{14}$$

while the network error is given as

$$e_i = Y^* - \hat{Y}^* \tag{15}$$

In this paper, we used the symmetric saturated linear transfer function,

$$f(x) = \begin{cases} -1, & x < -1 \\ x, & -1 \leq x \leq 1 \\ 1, & x > 1 \end{cases} \tag{16}$$

2.4 Data Source

The data used in this study were annual data on VVF patients admitted in Maryam Abacha Women and Children Hospital Sokoto. All input variables were standardized, that is, converting them to the range (0, 1) before feeding them into the network. This is to avoid the application of extremely small weighting factors in the case of large input values.

Similarly, the output values are "destandardized" to provide meaningful results since all values leaving the network are automatically output in a standardized format. This is done by simply reversing the standardization algorithm used on the input nodes.

Data processing and analysis were performed using Microsoft Windows versions of Microsoft Excel for the *Linear Regression* part of the analysis, while a neural code was written for the analysis of the using MATLAB R2009a

2.5 Model Comparison

The accuracy of predictions of *Linear Regression (LR) analysis or* statistical neural network *(SNN)* on the cause of VVF model was determined by Regression square, (R^2), mean square error (MSE), t-test statistics (t), F-test statistics (f) and P-value (p). These statistical methods were used in making comparisons.

3.0 Result and Discussions

The comparison of accuracy values of multiple regression models and neural network model are presented in Table 1 and Table 2 on Obstructed Labour and misuse of Instrument. In all cases, in comparison to traditional methods, ANN has a higher R square of 0.8 and 0.54 which is a better result, lower MSE 2011 and 4579.6 which is also a better result.

ADMISSION ON OBSTRUCTED

state after 1000 epochs

N = 10	K = 2	TANSIG_HN = 10
MEAN_err = 8.0932	VAR_err = 3.3944e+003	beta = -4.7852
alpha = 305.7219	T_test = 1.5652	SSE = 1.6088e+004
MSE = 2.0110e+003	F_cal = 31.9078	R2 = 0.7995
AIC = 2.4000e+003	SIC = 2.5498e+003	NIC = 1.9691e+003
NAR2 = 0.7745	p = 0.0760	

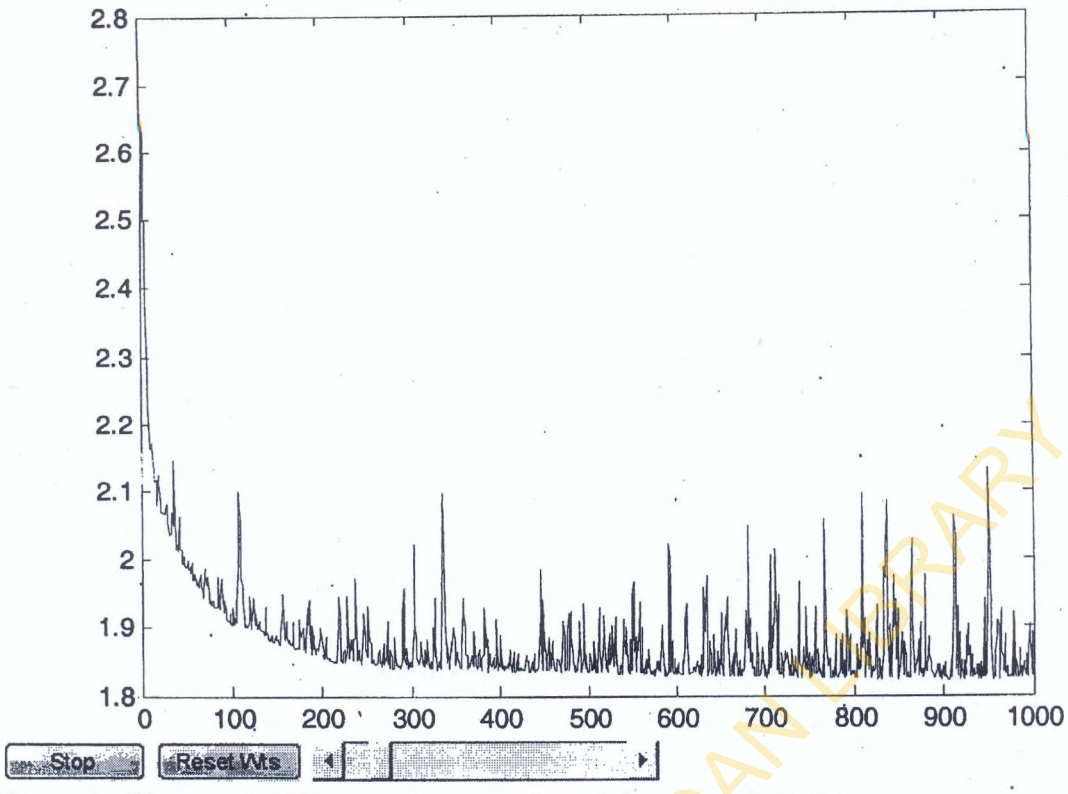

Figure 1 The graph of Error of propagation of the network training

Table1: The comparison of accuracy values of multiple regression models and neural network model in OBSTRUCTED LABOUR

	LR	ANN
R^2	0.46	0.8
MSE	5439.55	2011
t	2.6	1.57
F	6.75	31.91
p	0.03	0.08

Table2: The comparison of accuracy values of multiple regression models and neural network model in MISUSE OF INSTRUMENT

	LR	ANN
R^2	0.45	0.54
MSE	5269.99	4579.6
t	2.69	7.2
F	7.23	9.54
p	0.03	0.00003

Table 3: Anova 0n Regression Analysis On Obstructed Labour

r^2	0.458	n	10
r	-0.677	k	1
Std. Error	73.753	Dep. Var.	**ADMISSION**

ANOVA table

Source	*SS*	*df*	*MS*	*F*	*p-value*
Regression	36,738.0147	1	36,738.0147	6.75	.0317
Residual	43,516.3853	8	5,439.5482		
Total	**80,254.4000**	**9**			

Regression output

variables	*Coefficients*	*std. error*	*t (df=8)*	*p-value*	*confidence interval 95% lower*	*95% upper*
Intercept	305.7219	33.9519	9.005	.0000	227.4286	384.0152
OBSTRUCTED	-4.7852	1.8413	-2.599	.0317	-9.0313	-0.5392

LABOUR

Table 4: Anova for MISUSE OF INSTRUMENT

Regression Analysis

r^2	0.475	n	10
r	-0.689	k	1
Std. Error	72.595	Dep. Var.	**ADMISSION**

Source	*SS*	*df*	*MS*	*F*	*p-value*
Regression	38,094.4856	1	38,094.4856	7.23	.0276
Residual	42,159.9144	8	5,269.9893		
Total	**80,254.4000**	**9**			