
APPLICATION OF LOGISTIC REGRESSION MODELS FOR THE EVALUATION OF CHOLERA OUTBREAK IN ADAMAWA STATE NIGERIA

Bitrus Teri Musa & Olanrewaju Samuel Olayemi

Department of Statistics, University of Abuja, Abuja Nigeria

Corresponding Author email: olanrewaju.samuel@uniabuja.edu.ng

ABSTRACT: *Cholera outbreak occurred in four local government area of Adamawa state namely; Yola north, Yola south, Girei and Song between 11th may 2019 to 26th august 2019. WHO teams recorded 687 cholera patient whom received treatment at their various CTC with only 4 death as of the period of the outbreak. I explore cases of the disease outbreak, analyzed and estimate the parameters (demographic status and exploratory data) associated to the treatment outcome of the patients, identify the spread rate and targeted risk level using binary logistics regression. Our analysis has indicated that Yola North is the most affected area, majority of the patients are female and most of the respondent are children within the age group 1-14 years. The results depict that none of the demographic status was significantly associated with the mortality, similarly no significant association was observed for Culture status, Lab. Sample and Hospitalize status of the patients with mortality. However, the results of the association test between RDT status and mortality show that 2(0.3% of the total) who were RDT positive, were significantly associated with mortality. The binary logistic regression model estimations show that RDT, Culture, hospitalize status and LGA of the patients are the significant relationships at 10%, 10%, 5% and 10% levels respectively that determine the responses of patients to cholera treatment. Therefore, the findings depict that 99.4% of the patients “Alive” and 0.6% were “Dead”, this implies that the cholera treatment is very effective during the outbreak. Subsequently, it was deduced that the forecasting performance of the findings on both probit and logit regression model estimation from our analysis conform with the binary regression result.*

KEYWORD: cholera, logit, probit, epidemiology, RDT, binary regression

INTRODUCTION

Cholera was the first disease for which modern public health surveillance and reporting was carried out in an organized way. It is one of the three diseases currently reportable under the International Health Regulations (IHR) of 1969 while for Africa, the data began in 1970 with the acceptance of the International Health Regulations. According to those regulations, national health administrations should report the first cases of cholera on their

territory to WHO within 24 hours of their being informed. This applies to both indigenous cases of cholera as well as imported cases.

Cholera is an intestinal disease caused by the bacterium *Vibrio cholerae*, which colonizes the human intestine (Bertuzzo et al, 2010; Sack et al, 2004). *Vibrio cholera*, a curved Gram-negative bacillus belongs to the family, Vibrionaceae and shares some characteristics with the family, Enterobacteriaceae (Farmer, 2006). Cholera cases are confirmed through the isolation of *Vibrio cholerae* O1 or O139 from stools in any patient with diarrhea (WHO, 2004). Other serovars of *V. cholerae* are generally termed non-O1, non-O139 strains. They are non-choleraogenic, usually cause a milder form of gastroenteritis than O1 and O139, and are normally associated with sporadic cases and small outbreaks rather than with epidemics and pandemics (WHO, 2005). *Vibrio cholerae* O1 Eltor is the commonest strain in Nigeria (Utsalo et al, 1999; Opajobi et al, 2004; Usman et al, 2005; Shittu et al, 2010).

The disease cholera is characterized by diarrhea and severe dehydration. Only about 20 per cent of those infected develop acute, watery diarrhea (AWD), and of these, between 10–20 per cent develop severe watery diarrhea with vomiting. If people are not promptly and adequately treated, the loss of large amounts of fluid and salts through diarrhea and vomiting can lead to severe dehydration and death within hours. The case fatality rate (CFR) if untreated may reach 30–50 per cent. The typical presentation of cholera is a sudden onset of profuse, painless, watery stools, sometimes like rice-water, often accompanied by vomiting. Dehydration appears within 12–24 hours. The first 24 hours of cholera manifestation are the riskiest, and if the sufferer is not rehydrated, death can result. The main transmission mechanism for cholera is by drinking water or eating food contaminated by *Vibrio cholerae*, which enters the environment via feces (stools) from infected people.

In Nigeria, cholera is an endemic and seasonal disease, occurring annually mostly during the rainy season, with the first series of outbreaks reported between 1970 and 1990. Cholera epidemics in Nigeria can be traced back to 1961, but the first major epidemic, affecting 22,931 people with 2,945 deaths and a CFR of 12.8%, was reported in 1971. In 1991, another massive wave occurred, affecting 59,478 people with CFR of 12.9%, predominantly in the northern Nigerian.

Adamawa State has been frequently reporting cases of cholera outbreaks in recent time with high incidence of cases and deaths which are facilitated by numerous factors such as lack of access to safe drinking water, unhygienic environment, environmental disasters, literacy level, population congestion, and internal conflicts (Boko Haram farmers/herders clash) which may results to population displacement to Internally Displaced Persons (IDP) camps.

The Federal Government of Nigeria and partners will continue to work to reduce the impact of the current outbreak on affected communities and ensure that the outbreak is fully controlled. Therefore, in order to evaluate the disease outbreak by applying logistics regression models to determine association between the demographic status of patients and its exploratory data to the cholera treatment outcome, identify the spread rate, targeted risk level and provide a possible means to stop its reoccurrence, the discourse above formed a rational for this study.

Cholera Epidemiology in Nigeria

Cholera caused by *Vibrio cholera* continues to be a global threat to public health and a key indicator of lack of social development. Once common throughout the world, the infection is now largely confined to developing countries in the tropics and subtropics. It is endemic in Africa, parts of Asia, the Middle East, and South and Central America. In endemic areas, outbreaks usually occur when war or civil unrest disrupts public sanitation services. Natural disasters like earthquake, tsunami, volcanic eruptions, landslides and floods also contribute to outbreak by disrupting the normal balance of nature (Quadri, 2004). This creates many health problems, food and water supplies can become contaminated by parasites and bacteria when essential systems like those for water and sewage are destroyed. Developing countries are disproportionately affected because of their lack of resources, infrastructure and disaster preparedness systems (Sur, 2000). In newly affected areas, outbreaks may occur during any season and affect all ages equally. The organism normally lives in aquatic environments along the coast. People acquire its infection by consuming contaminated water, seafood, or other foods. Once infected, they excrete the bacteria in stool. Thus, the infection can spread rapidly, particularly in areas where human waste is untreated.

In Nigeria, the infection is endemic and outbreaks are not unusual. In the last quarter of 2009, it was speculated that more than 260 people died of cholera in four Northern states with over 96 people in Maidugari, Biu, Gwoza, Dikwa and Jere council areas of Bauchi state (Igomu, 2011). Most of the Northern states of Nigeria rely on hand dug wells and contaminated ponds as source of drinking water. Usually, the source of the contamination is other cholera patients when their untreated diarrhea discharge is allowed to get into water supplies (Igomu, 2011).

REVIEW OF LITERATURES

Ajoke O.A. et al, 2012 wrote on cholera epidemiology in Nigeria and explained that Cholera is an acute diarrhoeal infection caused by ingestion of food or water contaminated with the bacterium, *Vibrio cholera*. Choleraogenic *V. cholera* O1 and O139 are the only causative agents of the disease. They also state that the two most distinguishing epidemiologic features of the disease are its tendency to appear in explosive outbreaks and its predisposition to causing pandemics that may progressively affect many countries and spread into continents. On their research, they discover that despite efforts to control

cholera, the disease continues to occur as a major public health problem in many developing countries. Numerous studies over more than a century have made advances in the understanding of the disease and ways of treating patients, but the mechanism of emergence of new epidemic strains, and the ecosystem supporting regular epidemics, remain challenging to epidemiologists. In Nigeria, since the first appearance of epidemic cholera in 1972, intermittent outbreaks have been occurring. The later part of 2010 was marked with severe outbreak which started from the northern part of Nigeria, spreading to the other parts and involving approximately 3,000 cases and 781 deaths. Sporadic cases have also been reported. Although epidemiologic surveillance constitutes an important component of the public health response, publicly available surveillance data from Nigeria have been relatively limited to date. Based on existing relevant scientific literature on features of cholera, they conclude that this paper presents a synopsis of cholera epidemiology emphasizing the situation in Nigeria.

Apparently, Gregor C.L. and Auwol F.O. 2015 explore climate and socioeconomic influence on interannual variability of cholera in Nigeria. They deduce that, Cholera is one of the most important climate sensitive diseases in Nigeria that pose a threat to public health because of its fatality and endemic nature. Their study was aimed to investigate the influences of meteorological and socioeconomic factors on the spatiotemporal variability of cholera morbidity and mortality in Nigeria. The Step wise multiple regression and generalized additive models were fitted for individual states as well as for three groups of the states based on annual precipitation. Different meteorological variables were analyzed, taking into account socioeconomic factors that are potentially enhancing vulnerability (e.g. absolute poverty, adult literacy, access to pipe borne water). Their results quantify the influence of both climate and socioeconomic variables in explaining the spatial and temporal variability of the disease incidence and mortality. Regional importance of different factors is revealed, which will allow further insight into the disease dynamics. Additionally, cross validated models suggest a strong possibility of disease prediction, which will help authorities to put effective control measures in place which depend on prevention, and or efficient response.

David A.O. et al, 2012 did an assessment of the emergency response among health workers involved in the 2010 cholera outbreak in northern Nigeria which shows that the 2010 cholera outbreak in northern Nigeria affected over 40,000 people, with a case fatality rate (CFR) of $\geq 3.75\%$. They assessed the emergency response of health care workers (HCWs) involved in case management using a cross-sectional study with data collected through self-administered questionnaire. A total of 56 HCWs were interviewed. The mean age was 31 years ($SD \pm 8.16$ years). The majority of the HCWs (80%; $n = 45$) were aged 18–39 years. Most were community health extension workers (60%), and 3.6% ($n = 2$) were medical doctors. Many of the HCWs had less than 2 years of work experience (42%). Additionally, 82% of the respondents had <1 week of cholera emergency response training, and 50% of the HCWs managed >20 suspected cases of cholera per day. Although 78% of

HCWs reported the practice of universal safety precautions, 32% ($n = 18$) knew HCWs who developed symptoms of cholera during the epidemic, most of which was believed to be hospital acquired (78%). They also found that 77% ($n = 43$) of HCWs had no access to the required emergency response supplies. They conclude that Inadequate training, a lack of qualified HCWs and a limited supply of emergency response kits were reported.

RESEARCH METHODOLOGY

Study Area

This study is conducted at Adamawa which is in the northeastern Nigeria, with its capital at Yola. With four administrative divisions: Adamawa, Michika, Ganye, Mubi and Numan. Adamawa is one of the largest states of Nigeria located on the coordinates: 9°20'N 12°30'E and occupies total area of about 36,917 square kilometers (14,254 sq mi). it is bordered by the states of Borno to the northwest, Gombe to the west and Taraba to the southwest. Its eastern border forms the national eastern border with Cameroon. Topographically, it is a mountainous land crossed by the large river valleys – Benue, Gongola and Yedsarem. The valleys of the mount Cameroon, Mandara Mountains and Adamawa Plateau form part of the landscape. Adamawa states comprises of twenty-one local government areas with several languages spoken. The major occupation of the people is farming as reflected in their two notable vegetation zones, the sub-sudan and northern guinea savanna zones. Their cash crops are cotton and groundnut while food crops include maize, yam, cassava, guinea corn, millet and rice. The village community living on the banks of the rivers engage in fishing while the Fulanis are cattle rearers. The state has a network of roads linking all parts of the country. Adamawa state has an estimated population of about 3,178,950. Religious group in the state are mostly Christian and Muslims.

Study Design

Cholera outbreak in Adamawa was identified to the state ministry of health by WHO in the following locations: Yola North, Yola South, Girei and Song with the support of Nigeria Centre for Disease Control, NGOs, INGOs and other stakeholders in Adamawa State. Cases are detected through household visit by WHO during the outbreak, hospital case search, Community case search, patient and health care workers' interview and health records review. Cases are mapped out according to their LGAs and there is micro planning for the outbreak response activities.

Population of the Study

The study populations were patients with cholera suspected cases which are either hospitalized or not whom provide lab sample for both RDT and culture test with two possible outcomes death or alive.

Data Collection and Management

For the purpose of data collection in this research, details of all patients suspected with cholera cases identified at their various household during onset of the outbreak in Yola North, Yola South, Girei and Song whom are referred to the CTC by the WHO teams were computed on different section in Microsoft excel to determine spread rate, affected and targeted risk level among them.

Data Analysis

Data were entered into the computer and analyzed using IBM Statistical Package for Social Sciences (SPSS) version 23. Frequency distribution tables, charts and graphs were generated from variables while cross tabulation and test statistics done where applicable. Other methods of analysis are subsequently discussed.

The Logit and Logistic Transformation

In multiple regression, a mathematical model of a set of explanatory variables is used to predict the mean of a continuous dependent variable. In logistic regression, a mathematical model of a set of explanatory variables is used to predict a *logit* transformation of the dependent variable. Suppose the numerical values of 0 and 1 are assigned to the two outcomes of a binary variable. Often, the 0 represents a negative response and the 1 represents a positive response. The mean of this variable will be the proportion of positive responses. If p is the proportion of observations with an outcome of 1, then $1-p$ is the probability of an outcome of 0. The ratio $p/(1-p)$ is called the *odds* and the *logit* is the logarithm of the odds, or just *log odds*. Mathematically, the logit transformation is written

$$l = \text{logit } p = \ln \frac{p}{1-p}$$

The *logistic* transformation is the inverse of the logit transformation. It is written

$$p = \text{logistic}(l) = \frac{e^l}{1 + e^l}$$

The Logistic Regression and Logit Model

In logistic regression, a categorical dependent variable Y having G (usually $G = 2$) unique values is regressed on a set of p independent variables X_1, X_2, \dots, X_p

Let

$$X = (X_1, X_2, \dots, X_p)$$

$$B_g = \begin{pmatrix} \beta_{g1} \\ \vdots \\ \beta_{gp} \end{pmatrix}$$

The logistic regression model is given by the G equations

$$\ln \left(\frac{P_g}{P_1} \right) = \ln \left(\frac{P_g}{P_1} \right) + \beta_{g1}X_1 + \beta_{g2}X_2 + \dots + \beta_{gp}X_p$$

$$= \ln\left(\frac{P_g}{P_1}\right) + XB_g$$

Here, p_g is the probability that an individual with values X_1, X_2, \dots, X_p is in outcome g . That is,

$$p_g = \Pr(Y = g|X)$$

Usually $X_1 \equiv 1$ (that is, an intercept is included), but this is not necessary.

The quantities P_1, P_2, \dots, P_g represent the prior probabilities of outcome membership. If these prior probabilities are assumed equal, then the term $\ln(P_g/P_1)$ becomes zero and drops out. If the priors are not assumed equal, they change the values of the intercepts in the logistic regression equation. Outcome one is called the *reference value*. The regression coefficients $\beta_{11}, \beta_{12}, \dots, \beta_{1p}$ for the reference value are set to zero. The choice of the reference value is arbitrary. Usually, it is the most frequent value or a control outcome to which the other outcomes are to be compared. This leaves $G-1$ logistic regression equations in the logistic model.

The β 's are population regression coefficients that are to be estimated from the data. Their estimates are represented by b 's. The β 's represents unknown parameters to be estimated, while the b 's are their estimates.

These equations are linear in the logits of p . However, in terms of the probabilities, they are nonlinear. The corresponding nonlinear equations are

$$P_g = \text{Prob}(Y = g|X) = \frac{e^{XB_g}}{1 + e^{XB_2} + e^{XB_3} + \dots + e^{XB_G}}$$

since $e^{XB_1} = 1$ because all of its regression coefficients are zero.

A note on the names of the models. Often, all of these models are referred to as *logistic regression models*.

However, when the independent variables are coded as ANOVA type models, they are sometimes called *logit models*. A note about the interpretation of e^{XB} may be useful. Using the fact that $e^{XB} = (e^a)(e^b)$, (e^{XB}) may be re expressed as follows

$$\begin{aligned} e^{XB} &= e^{\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p} \\ &= e^{\beta_1 X_1} e^{\beta_2 X_2} \dots e^{\beta_p X_p} \end{aligned}$$

This shows that the final value is the product of its individual terms. (NCCS)

Binary Logistic Regression

Binary logistic regression estimates the probability that a characteristic is present (e.g. estimate probability of "success") given the values of explanatory variables, in this case a single categorical variable; $\pi = \Pr(Y = 1|X = x)$.

Variables:

- Let Y be a binary response variable
- $Y_i = 1$ if the trait is present in observation (person, unit, etc...) i
- $Y_i = 0$ if the trait is NOT present in observation i

- $X = (X_1, X_2, \dots, X_k)$ be a set of explanatory variables which can be discrete, continuous, or a combination. x_i is the observed value of the explanatory variables for observation i . In this section of the notes, we focus on a single variable X .

$$\pi_i = \Pr(Y_i = 1|X_i = x_i) = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}$$

Or,

$$\begin{aligned} \text{logit}(\pi_i) &= \log\left(\frac{\pi_i}{1 - \pi_i}\right) \\ &= \beta_0 + \beta_1 x_i \\ &= \beta_0 + \beta_1 x_{i1} + \beta_k x_{ik} \end{aligned}$$

Probit Analysis

Probit analysis is a method of analyzing the relationship between a stimulus (dose) and the quantal (all or nothing) response. Quantitative responses are almost always preferred, but in many situations, they are not practical. In these cases, it is only possible to determine if a certain response (such as death) has occurred.

The Probit Model assumes that the percent response is related to the log dose as the cumulative normal distribution. That is, the log doses may be used as variables to read the percent dying from the cumulative normal.

Using the normal distribution, rather than other probability distributions, influences the predicted response rate at the high and low ends of possible doses, but has little influence near the middle. Hence, much of the comparison of different drugs is done using response rates of fifty percent. The probit model may be expressed mathematically as follows:

$$P = \alpha + \beta[\log_{10}(\text{Dose})]$$

where P is five plus the inverse normal transform of the response rate (called the Probit). The five is added to reduce the possibility of negative probits, a situation that caused confusion when solving the problem by hand. (NCCS Statistical Software)

Likelihood Ratio

The *Likelihood Ratio* test statistic is -2 times the difference between the log likelihoods of two models, one of which is a subset of the other. The distribution of the LR statistic is closely approximated by the chi-square distribution for large sample sizes. The degrees of freedom (DF) of the approximating chi-square distribution is equal to the difference in the number of regression coefficients in the two models. The test is named as a ratio rather than a difference since the difference between two log likelihoods is equal to the log of the ratio of the two likelihoods. That is, if L_{full} is the log likelihood of the full model and L_{subset} is the log likelihood of a subset of the full model, the likelihood ratio is defined as

$$\begin{aligned} LR &= -2[L_{\text{subset}} - L_{\text{full}}] \\ &= -2\left[\ln\left(\frac{l_{\text{subset}}}{l_{\text{full}}}\right)\right] \end{aligned}$$

Note that the -2 adjusts LR so the chi-square distribution can be used to approximate its distribution.

The likelihood ratio test is the test of choice in logistic regression. Various simulation studies have shown that it is more accurate than the Wald test in situations with small to moderate sample sizes. In large samples, it performs about the same. Unfortunately, the likelihood ratio test requires more calculations than the Wald test, since it requires that two maximum-likelihood models must be fit.

4DATA PRESENTATION AND ANALYSIS

Data Presentation

The data utilized in this research were secondary sourced from Adamawa State Ministry of Health by WHO. Data comprised of 687 cholera patients receiving treatments.

Data Analysis

Demographic Data of Respondents

The demographic information of the cholera patients (Table 4.1) reveals that 42.6% (293) of the patients were men and 57.4% (394) of them were women receiving treatment. The age pattern of the patients reveals that majority of them were between 1-4years (27.2%) and 5-14years (20.8) while 14.0%, 13.0%, 12.1% 9.3% and 3.6% of them were between 15-24years, 25-34years, ≥ 45 years, 35-44years and less than 1year respectively. The results further show that majority of receiving cholera treatment were from Y-North local government area (LGA) (61.3%) while 25.8%, 12.8% and 0.1% of them were from Girei, Y-South and Song LGAs respectively.

Table 4.1: Demographic Data of Cholera Patients

Variables	Frequency (%)
Age	
Less than 1 Year	25 (3.6)
1-4 Year(s)	187 (27.2)
5-14 Years	143 (20.8)
15-24 Years	96 (14.0)
25-34 Years	89 (13.0)
35-44 Years	64 (9.3)
45 and above Years	83 (12.1)
Gender	
Male	293 (42.6)
Female	394 (57.4)
LGA	
Girei	177 (25.8)
Song	1 (0.1)
Y-North	421 (61.3)
Y-South	88 (12.8)

Source: Field Survey, 2019

Furthermore, the treatment outcomes of the patients receiving cholera treatments as at under examination were revealed in Figure 4.0. The results depict that 99.4% of the patients “Alive” and 0.6% were “Dead”, hence this implies that the cholera treatment is very effective.

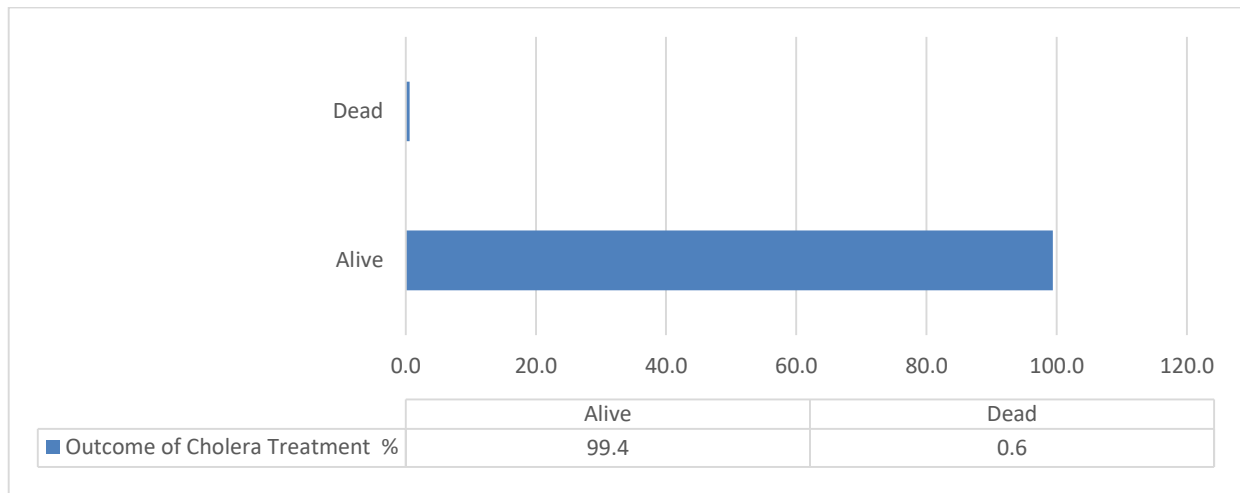


Figure 4.0: Outcome of Cholera Treatment

Demographic status, Cholera Treatment and Mortality

Table 4.2 and Figure 4.1 to 4.3 reveal the associations between demographic status of patients and cholera treatment outcome. The results show the mortality distribution of patients receiving treatment across their demographic status: a total of 4 (0.6% of the total) patients were recorded dead which 2(0.3%), 1(0.1%) and 1(0.1%) were between 1-4year(s), 25-34years and ≥ 45 years age respectively; 3(0.4%) were male and 1(0.2%) were female; 1(0.1%), 2(0.3%) and 1(0.1%) were from the Girei, Y-North and Y-South respectively. However, none of the demographic status is statistically significant association with the mortality.

Table 4.2: Association between Demographic Status and Cholera Treatment Outcome among Patients (Cross-tabulation)

Variables	Alive (%)	Dead (%)	X ² (P-value)
Age			
Less than 1 Year	25 (3.6)	0	3.694 (0.718) DF=6
1-4 Year(s)	185 (26.9)	2 (0.3)	
5-14 Years	143 (20.8)	0	
15-24 Years	96 (14.0)	0	
25-34 Years	88 (12.8)	1 (0.1)	
35-44 Years	64 (9.3)	0	
45 and above Years	82 (11.9)	1 (0.1)	
Gender			
Male	290 (42.2)	3 (0.4)	1.722 (0.318) DF=1
Female	393 (57.2)	1 (0.2)	
LGA			
Girei	176 (25.6)	1 (0.1)	0.557 (0.906) DF=3
Song	1 (0.1)	0	
Y-North	419 (61.0)	2 (0.3)	
Y-South	87 (12.7)	1 (0.1)	

Source: Researchers' computations Mortality Versus RDT Test, Culture Test, Laboratory Sample and Hospitalize status of Patients

Table 4.3 depicts the associations between treatment outcome and RDT test, culture test, laboratory sample & hospitalize status of patients. The mortality results show that 2(0.3% of the total) were RDT positive that were significantly associated with mortality. However, 2(0.3%) and 2(0.3%) had laboratory sample test and had no laboratory sample test respectively that are not significantly associated with mortality. Similarly, 2(0.3%) and 2(0.3%) were hospitalized and not hospitalized respectively that are not significantly associated with mortality.

Table 4.3: Association between RDT Test, Culture Test, Laboratory Sample and Hospitalize status of Patients and their Treatment Outcome

Variables	Alive (%)	Dead (%)	X ² (P-value)
RDT Test			
Positive	210 (30.9)	2 (0.3)	1.227 (0.047) DF=3
Negative	129 (19.0)	0	
Inconclusive	1 (0.1)	0	
NA	336 (48.9)	2 (0.3)	
Culture Test			
Positive	79 (18.9)	0	0.476 (0.788) DF=2
Negative	1 (0.2)	0	
NA	336 (80.4)	4 (1.0)	
Laboratory Sample			
Yes	340 (50.0)	2 (0.3)	0.567 0.991) DF=1
No	336 (49.4)	2 (0.3)	
Hospitalize Status			
Yes	282 (41.0)	2 (0.3)	0.124 (0.724) DF=1
No	401 (58.4)	2 (0.3)	

Source: Researchers' computations

Binary Logistic Regression Model Estimation

The binary logistic regression estimation results for cholera treatment outcome are presented in Table 4.4 to Table 4.8. The results in Table 4.4 show the rates of the two treatment outcomes ($416/418 = 99.5\%$ are alive and 0.5% were dead). the best strategy is to predict for every case, that the subject will be alive. Using that strategy well shall be correct 99.5% of the time.

Table 4.4: Classification Table

	Observed	Predicted		
		Outcome		Percentage Correct
		Alive	Dead	
Step 0	Outcome Alive	416	0	100.0
	Dead	2	0	.0
Overall Percentage				99.5

Source: IBM SPSS 23 **Note:** Constant is included in the model & the cut value is .500

Under the variables in the Equation (Table 4.5) we see that the constant-only model is $\ln(odds) = -5.338$. This implies that the predicted odds of 'dead' is 0.005.

Table 4.5: Variables in the Equation

	β	S.E.	Wald	df	Sig.	Exp(β)
Step 0 Constant	-5.338	.709	56.706	1	.000	.005

Source: IBM SPSS 23

Table 4.6 presents the Omnibus Tests of Coefficients results for the predictors. The results give a Chi-square of 5.920 on 6 degree of freedom, significant at 5% level of significance. Hence, it implies adding the predictors (gender, age-group, RDT status, culture status, hospitalize status and LGA) have significantly increased our ability to predict the cholera outcomes (dead).

Table 4.6: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	5.920	6	.049
	Block	5.920	6	.049
	Model	5.920	6	.049

Source: IBM SPSS 23

Table 4.7 presents the summary of logistic regression model. The -2 Log likelihood statistic is 19.440 (this measure how poorly the model predicts the cholera treatment outcomes; the smaller it is the better). The Cox & Snell R^2 and Nagelkerke R^2 depict the amount of variations (1.4% and 23.9% respectively) in cholera treatment outcomes that are explained by the predictors.

Table 4.7: Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	19.440 ^a	.014	.239

Source: IBM SPSS 23

Note: ^a Estimation terminated at iteration number 20 because maximum iterations has been reached.

Table 4.8 presents the binary regression estimations, it shows the relationship between the dependent (treatment outcome) and independent variables; gender (sex), age-group (AG), RDT, Culture, hospitalize status (HOS) and LGA. The RDT, Culture, hospitalize status and LGA of the patients are the significant relationships at 10%, 10%, 5% and 10% levels respectively. Hence, RDT, Culture, hospitalize status and LGA of the patients predicts the treatment outcome (dead). The model estimations show that the logistic regression equation is

$$\ln(ODDS) = -29.09 + 1.42LGA - 0.26HOS - 0.3Culture - 0.48RDT + 0.14AG - 16.81Sex$$

4.1

Table 4.8: Binary Logistic Regression Estimation

		β	S.E.	Wald	Sig.	Exp(β)
Step 1	Sex	-16.814	2372.629	.000	.994	.000
	AG	.140	.428	.107	.744	1.150
	RDT*	-.477	1.167	.167	.083	.621
	Culture*	-.301	1.357	.049	.075	.740
	HOS**	-.258	1.256	.042	.037	.772
	LGA*	1.416	.818	2.997	.083	4.120
	Constant	-29.092	7650.204	.000	.997	.000

Source: IBM SPSS23**Note:** * and ** denote significant at 10% and 5% respectively**Probit Regression Model Estimation**

Table 4.10 depicts the Omnibus test results. The results give a likelihood ratio chi-square of 10.977 with degree of freedom 15, significant at 5% level of significance. Hence, similar to the Omnibus Test result in Table 4.6, Table 4.9 results imply that adding the predictors (gender, age-group, RDT status, culture status, hospitalize status and LGA) have significant increase in the ability to predict the cholera outcomes (dead).

Table 4.10: Omnibus Test (Probit)

Likelihood Ratio Chi-Square	Df	Sig.
10.977	15	.045

Source: IBM SPSS 23

Dependent Variable: Outcome; Model: (Intercept), AG, RDT, Culture, LGA, sex, HOS Compares the fitted model against the intercept-only model

Table 4.11 presents the results of models' effects test. The shows that RDT, Culture, LGA and HOS are the predictor whose effects are significant. Table 4.12 presents the probit regression estimation of the parameters

Table 4.11: Tests of Model Effects

Source	Type III		
	Wald Chi-Square	Df	Sig.
(Intercept)	18.160	1	1.000
AG	2.246	6	.896
RDT	4.754	2	.079
Culture	2.311	1	.050
LGA	3.195	3	.048
Sex	.000	1	.997
HOS	1.214	1	.045

Source: IBM SPSS 23

Dependent Variable: Outcome; Model: (Intercept), AG, RDT, Culture, LGA, sex, LS, HOS

Table 4.12: Parameter Estimation

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(β)
			Lower	Upper	Wald Chi-Square	Df	Sig.	
(Intercept)	16.273	3224.2559	-11359.859	11392.406	2.320	1	.998	11678991.082
[AG=1.00]	.421	1.1520	-2.439	2.281	.000	1	1.000	1.523
[AG=2.00]	-4.024	.9846	-5.652	5.604	.001	1	.999	.018
[AG=3.00]	.336	1.5731	-4.614	7.286	.000	1	1.000	1.399
[AG=4.00]	.193	.2888	-73.832	76.219	.001	1	.988	1.213
[AG=5.00]	-5.403	.9847	-4.031	5.225	.002	1	.998	.005
[AG=6.00]	.037	5.5754	-1.539	8.612	.000	1	1.000	1.037
[RDT=1.00]**	-4.876	6.1880	-5.186	5.434	5.147	1	.023	.008
[RDT=2.00]**	-4.851	8.5058	-50.122	54.420	3.001	1	.045	.008
[RDT=3.00]**	-14.916	8.8299	-73.994	77.161	5.023	1	.042	3.327E-7
[Culture=1.00]*	-4.957	4.9495	-2.935	5.848	6.071	1	.031	142.115
[LGA=1.00]	4.916	2.2891	-60.887	62.055	.001	1	.989	.007
[LGA=2.00]	5.540	7.4070	-56.647	54.567	.020	1	.892	1.783E-7
[LGA=3.00]**	15.302	.2892	-61.273	62.669	5.980	1	.019	.005
Sex	5.499	1374.7827	-2689.025	2700.024	.210	1	.997	244.512
HOS**	-15.270	1.8737	-3.634	3.095	4.310	1	.038	.005
(Scale)	1 ^c							

Dependent Variable: Outcome; Model: (Intercept), AG, RDT, Culture, LGA, sex, HOS ** denote significant at 10% **Source: IBM SPSS 23**

Table 4.12 presents the probit regression estimations, it shows the relationship between the dependent (treatment outcome; Dead) and independent variables; gender (sex), age-group

(AG), RDT, Culture, hospitalize status (HOS) and LGA. The RDT, Culture, hospitalize status of the patients have significant negative effect on the treatment outcome (Dead) at 5% level while the LGA of the patients have significant positive effect on the treatment outcome (Dead) at 5% level. This implies that cholera patients that are taken to hospital, and undergo necessary test such as RDT and Culture have more likelihood of being alive, while patients from Y-North LGA are more likely to die. Hence these findings conform to the binary results presented in Table 4.8

Logit Regression Model Estimation

Table 4.14 depicts the Omnibus test results. The results give a likelihood ratio chi-square of 10.854 with degree of freedom 15, significant at 5% level of significance. Hence, similar to others' Omnibus Test results (see Table 4.6 and Table 4.10). This also implies adding the predictors (gender, age-group, RDT status, culture status, hospitalize status and LGA) have significant increase in the ability to predict the cholera outcomes (dead).

Table 4.14: Omnibus Test (Logit)

Likelihood Ratio Chi-Square	Df	Sig.
10.854	15	.049

Dependent Variable: Outcome; Model: (Intercept), AG, RDT, Culture, LGA, sex, HOS

Source: IBM SPSS 23

Table 4.15 presents the results of models' effects test. It shows that RDT, Culture, LGA and HOS are the predictor whose effects are significant. Table 4.16 presents the logit regression estimation of the parameters

Table 4.15: Tests of Model Effects

Source	Type III		
	Wald Chi-Square	Df	Sig.
(Intercept)	.000	1	1.000
AG	2.333	6	.887
RDT	6.350	2	.037
Culture	4.057	1	.045
LGA	5.087	3	.022
Sex	.000	1	.996
HOS	7.304	1	.008

Source: IBM SPSS 23

Dependent Variable: Outcome; Model: (Intercept), AG, RDT, Culture, LGA, sex, HOS

Table 4.16 presents the logit regression estimations, it shows the relationship between the dependent (treatment outcome; Dead) and independent variables; gender (sex), age-group (AG), RDT, Culture, hospitalize status (HOS) and LGA.

Table 4.16: Parameter Estimation

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(β)
			Lower	Upper	Wald Chi-Square	Df	Sig.	
(Intercept)	53.682	17363.8955	-33978.928	34086.292	.000	1	.998	2059319.000
[AG=1.00]	1.966	3.8823	-62.322	63226.253	.000	1	1.000	7.140
[AG=2.00]	-15.799	8.0013	-16.817	16.220	.001	1	.999	1.376E-7
[AG=3.00]	1.555	6.2877	-21.174	21.283	.000	1	1.000	4.733
[AG=4.00]	.777	11.6614	-22.459	22.012	.001	1	.988	2.174
[AG=5.00]	-18.425	8.0013	-16.444	16.593	.000	1	.998	9.954E-9
[AG=6.00]	.207	12.3202	-24.767	24.180	.002	1	.999	1.230
[RDT=1.00]*	-5.195	8.1259	-1.380	1.989	7.147	1	.012	2.516E-7
[RDT=2.00]	-5.259	8.1058	-1.622	1.105	2.001	1	.055	2.362E-7
[RDT=3.00]	-18.007	1.3127	-2.977	2.964	.045	1	.752	1.415E-21
*[Culture=1.00]	-5.709	80.2231	-1.256	1.674	2.651	1	.064	6640511.436
[LGA=1.00]	17.310	9.8890	-18.189	18.568	.001	1	.999	3.035E-8
[LGA=2.00]	51.622	8.3544	-15.521	15.276	.002	1	.992	3.807E-23
[LGA=3.00]	17.816	9.8890	-18.695	18.063	9.980	1	.016	1.831E-8
Sex	18.811	4073.6034	-7965.305	8002.927	.010	1	.999	147765784.143
HOS**	-17.629	5.4171	-11.842	11.585	5.410	1	.028	2.208E-8
(Scale)	1 ^c							

Note: * ** denote significant at 5% level **Source: IBM SPSS 23**

Similarly, to probit regression results the RDT, Culture, hospitalize status of the patients have significant negative effect on the treatment outcome (Dead) at 5% level while the LGA of the patients have significant positive effect on the treatment outcome (Dead) at 5% level. This implies that cholera patients that are taken to hospital, and undergo necessary

test such as RDT and Culture have more likelihood of being alive, while patients from Y-North LGA are more likely to die. Hence these findings conform to the binary results presented in Table 4.8.

Summary

The study is aimed at providing an extensive analysis and explore cholera outbreak in Adamawa State using logistics regression models. It also sought to explore the association between demographic status of patients and cholera treatment outcomes. Additionally, it addressed the associations of some of the exploratory data such as RDT test, Culture test, Lab. Sample and Hospitalize status of the cholera patients with the treatment outcomes.

CONCLUSION

Majority of the respondents analyzed are within the children age group 1-14years [1-4years (27.2%) and 5-14years (20.8)]. It also shows that majority of the patients were women (57.4%) and were from Y-North local government area (61.3%). In terms of gender, there were more females than males closely to ratio 3:2 (see Table 4.1). This was identical to study done by Noelle (2013) on gender and vulnerability to cholera in Sierra Leone. Additionally, the findings depict that 99.4% of the patients “Alive” and 0.6% were “Dead”, hence this implies that the cholera treatment is very effective (see Figure 4.1).

Furthermore, the associations between demographic status of patients and cholera treatment outcome were tested, the results depict that none of the demographic status was statistically significant association with the mortality (see Table 4.2). Similarly, no significant association was observed for Culture status, Lab. Sample and Hospitalize status of the patients with mortality. However, the results of the association test between RDT status and mortality show that 2(0.3% of the total) who were RDT positive, were significantly associated with mortality.

Moreover, the binary logistic regression model estimations show that RDT, Culture, hospitalize status and LGA of the patients are the significant relationships at 10%, 10%, 5% and 10% levels respectively that determine the responses of patients to cholera treatment. Therefore, RDT, Culture, hospitalize status and LGA of the patients predicts the treatment outcome. Hence, we deduce that the forecasting performance of the findings on both probit and logit regression model estimation from our analysis conform the binary regression result.

Recommendations

In consonance with the findings of this research, the following recommendations are made:

- i. Y-North local government area should be given more attention in terms of proper hygiene as many people in the area are prone to cholera;
- ii. Children between the age 1-14years should be given proper hygienic cares;
- iii. More supports should be given to the health facilities in order to perform more efficient in dealing with cholera outbreaks.

- iv. Attention should also be given to areas bordering the affected location

References

- Adamawa State Ministry of Health Cholera Situation Report, 2019
- Ajoke Olutola Adagbada, Solayide Abosedo Adesida, Francisca Obiageri Nwaokorie, Mary-Theresa Niemogha, Akitoye Olusegun Coker (2012) “Cholera Epidemiology in Nigeria: an overview” The Pan African Medical Journal - ISSN 1937-8688.
- Bertuzzo, E., Casagrandi, R., Gatto, M., Rodriguez-Iturbe, I., Rinaldo, A., (2010). On spatially explicit models of cholera epidemics. *J. R. Soc. Interf.* 7 (43), 321e333.
- David A. Oladele, Kolawole S. Oyedele, Mary-Theresa Niemogha, Francisca Nwaokorie, Moses Bamidele, Adesola Z. Musa, Adeniyi K. Adeneye, Tajudeen A. Bamidele, Michael Ochoga, Kehinde A. Akinsinde, Bartholomew I. Brai, Emmanuel A. Omonigbehin, Toun W. Fesobi, Stella I. Smith, Innocent A. Ujah. (2012) “An assessment of the emergency response among health workers involved in the 2010 cholera outbreak in northern Nigeria” *Journal of Infection and Public Health* 5, 346—353
- David W. Hosmer, Jr., Stanley Lemeshow, Rodney X. Sturdivant. (2013). *Applied Logistic Regression / 3rd Edition* ISBN 978-0-470-58247-3
- Gregor C. Leckebusch, Auwal F. Abdussalam. (2015) “Climate and socio-economic influences on interannual variability of cholera in Nigeria” *Health & Place* 34. Pg. 107–117.
- Igomu T. (2011) Cholera epidemic; 23 August 2011.
- Opajobi SO, Kandakai OYT, Mawak JD, Olukemi MA, Bello CSS (2004) *Vibrio cholerae* 01 infections in Jos, Nigeria. *African J Clin Exp Biol.*; 5(3):260-264.
- Quadri F.(2005) Enterotoxigenic *Escherichia coli* and *Vibrio cholerae* diarrhea, Bangladesh, 2004. *Emerg Infect Dis.* Jul;11(7):1104-7.
- Sack, D.A., Sack, R.B., Nair, G.B., Siddique, A.K., (2004). Cholera. *Lancet* 363 (9404), 223e233.
- Shittu OB, Akpan TO, Popoola S, Oyedepo JA, Ogunshola EO. (2010) Epidemiological features of a GIS-supported investigation of cholera outbreak in Abeokuta, Nigeria. *J Public Health Epidemiol.*2(7):152–162
- Sur D. (2000). Severe cholera outbreak following floods in a northern district of West Bengal. *Indian J Med Res.* Nov;112:178-82.
- UNICEF Situation Report (NIGERIA) Geneva: (2010). Cholera and floods Sitrep #5 covering 23 September-4 October
- Usman A, Sarkinfada F, Mufunda J, Nyarango P, Mansur K, Daiyabu TM. (2005) Recurrent cholera epidemics in Kano-Northern Nigeria. *Cent Afr J Med.* Mar-Apr;51(3-4):34-8.
- Utsalo SJ, Eko FO, Umoh F, Asindi AA. (1999) Faecal excretion of *Vibrio cholerae* during convalescence of cholera patients in Calabar, Nigeria. *Eur J Epidemiol.* Apr;15(4):379-81.

World Health Organization. Geneva: (2004). Communicable Disease Toolkit for Sierra Leone (2004): Communicable Disease Profile WHO/CDS/2004

World Health Organization/Food and Agriculture Organization (2005). Microbiological Risk Assessment Series; 2005. Risk assessment of cholerae *Vibrio cholerae* 01 and 0139 in warm-water shrimp in international trade.

http://www.unesco.org/science/disaster/about_disaster.shtml Accessed 29 June 2012.

<https://www.afro.who.int/news/who-partners-ncdc-effective-containment-cholera-outbreaks>

Cholera outbreak: assessing the outbreak response and improving preparedness:

<http://www.who.int/cholera/publications/OutbreakAssessment/en/index.html>

Ensuring food safety in the aftermath of natural disasters

Appendix

Gender * Outcome Crosstabulation

Count

		Outcome		Total
		Alive	Dead	
Gender	Male	290	3	293
	Female	393	1	394
Total		683	4	687

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.722 ^a	1	.189		
Continuity Correction ^b	.648	1	.421		
Likelihood Ratio	1.736	1	.188		
Fisher's Exact Test				.318	.210
Linear-by-Linear Association	1.719	1	.190		
N of Valid Cases	687				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 1.71.

b. Computed only for a 2x2 table

Age Group * Outcome Cross tabulation

Count

		Outcome		Total
		Alive	Dead	
Age Group	less than Year	25	0	25
	1-4 Years	185	2	187
	5-14 Years	143	0	143
	15-24 Years	96	0	96
	25-34 Years	88	1	89
	35-44 Years	64	0	64
	45 and Above	82	1	83
Total		683	4	687

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	3.694 ^a	6	.718
Likelihood Ratio	5.223	6	.516
Linear-by-Linear Association	.042	1	.837
N of Valid Cases	687		

a. 7 cells (50.0%) have expected count less than 5. The minimum expected count is .15.

RDT Test * Outcome Cross tabulation

Count

		Outcome		Total
		Alive	Dead	
RDT Test	Positive	210	2	212
	Negative	129	0	129
	Inconclusive	1	0	1
	NA	336	2	338
Total		676	4	680

Chi-Square Tests

	Value	df	Asymptotic Significance (2- sided)
Pearson Chi-Square	1.227 ^a	3	.047
Likelihood Ratio	1.920	3	.059
Linear-by-Linear Association	.074	1	.076
N of Valid Cases	680		

a. 5 cells (62.5%) have expected count less than 5. The minimum expected count is .01.

Culture Test * Outcome Cross tabulation

		Outcome		Total
		Alive	Dead	
Culture Test	Positive	79	0	79
	Negative	1	0	1
	NA	336	2	338
Total		416	2	418

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	.476 ^a	2	.788
Likelihood Ratio	.852	2	.653
Linear-by-Linear Association	.473	1	.492
N of Valid Cases	418		

a. 4 cells (66.7%) have expected count less than 5. The minimum expected count is .00.

Laboratory Sample * Outcome Cross tabulation

Count

		Outcome		Total
		Alive	Dead	
Laboratory Sample	Yes	340	2	342
	No	336	2	338
Total		676	4	680

Hospitalized * Outcome Cross tabulation

Count

		Outcome		Total
		Alive	Dead	
Hospitalized	Yes	282	2	284
	No	401	2	403
Total		683	4	687

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.124 ^a	1	.724	1.000	.548
Continuity Correction ^b	.000	1	1.000		
Likelihood Ratio	.123	1	.726		
Fisher's Exact Test					
Linear-by-Linear Association	.124	1	.724		
N of Valid Cases	687				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 1.65.

b. Computed only for a 2x2 table