

# Analysis on Reported Cases of HIV at Ekiti State University Teaching Hospital, Ado-Ekiti, South-Western Nigeria

Kayode S. O. Ibikunle and Isaac O. Ajao

*Department of Mathematics and Statistics, The Federal Polytechnic, Ado-Ekiti, Nigeria*  
[isaacoluwaseyiajao@gmail.com](mailto:isaacoluwaseyiajao@gmail.com)

## ABSTRACT

The nation Nigeria has been ranked the second among the nations of world with largest population of people with HIV. This information does not mean that every region and state in the country is not safe. Ekiti, a state in the south western part of the country has the lowest rate, which this study aims at establishing scientifically. Considering certain risk factors, such as age, gender and the local government area of the individuals tested, it is found out that none of the factors is significantly contributing to having or not having HIV in the state at 5% of significance. The parameter estimates obtained using the binary logistic regression are very low, the lack of fit test and the model test show that the factors are not good for modelling HIV cases in Ekiti state, thereby corroborating the fact that the state has lowest rate in Nigeria.

**Keyword:** HIV, Ekiti state, Nigeria, logistic regression

## 1 Introduction

Nigeria has the second largest HIV epidemic in the world (National Agency for the Control of AIDS (NACA), 2017). Although HIV prevalence among adults is much less (2.8%) than other sub-Saharan African countries such as South Africa (18.8%) and Zambia (11.5%), the size of Nigeria's population means 3.1 million people were living with HIV in 2017 (United Nations Joint Program on HIV/AIDS (UNAIDS), 2018) It is estimated that around two-thirds of new HIV infections in West and Central Africa in 2017 occurred in Nigeria. Together with South Africa and Uganda, the country accounts for around half of all new HIV infections in sub-Saharan Africa every year (UNAIDS, 2017). This is despite achieving a 5% reduction in new infections between 2010 and 2017 (UNAIDS, 2018). Unprotected heterosexual sex accounts for 80% of new HIV infections in Nigeria, with the majority of remaining HIV infections occurring in key affected populations such as sex workers (NACA, 2015). Six states in Nigeria account for 41% of people living with HIV, including Kaduna, Akwa-Ibom, Benue, Lagos, Oyo, and Kano (NACA, 2017). HIV prevalence is highest in Nigeria's southern states (known as the South South Zone), and stands at 5.5%. It is lowest in the southeast (the South East Zone) where there is a prevalence of 1.8%. There are higher rates of HIV in rural areas (4%) than in urban ones (3%) (NACA, 2015) Approximately 150,000 people died from AIDS-related illnesses in Nigeria in 2017 (UNAIDS, 2018). Since 2005, the reduction in the number of annual AIDS-related deaths has been minimal, indicative of the fact that only 33% of those with a positive diagnosis in Nigeria are accessing antiretroviral treatment (ART) (UNAIDS, 2018). According to the HIV/AIDS prevalence ranking by states published by National Agency for the Control of Aids (NACA) Ekiti,

a south western state in Nigeria has the lowest case of HIV (NACA, 2017). This research reveals this information about Ekiti state and establishes the reason behind the insignificant determinants of HIV in the state as shown in the results.



Fig. 1: Annual number of AIDS-related deaths in Nigeria (UNAIDS, 2018)

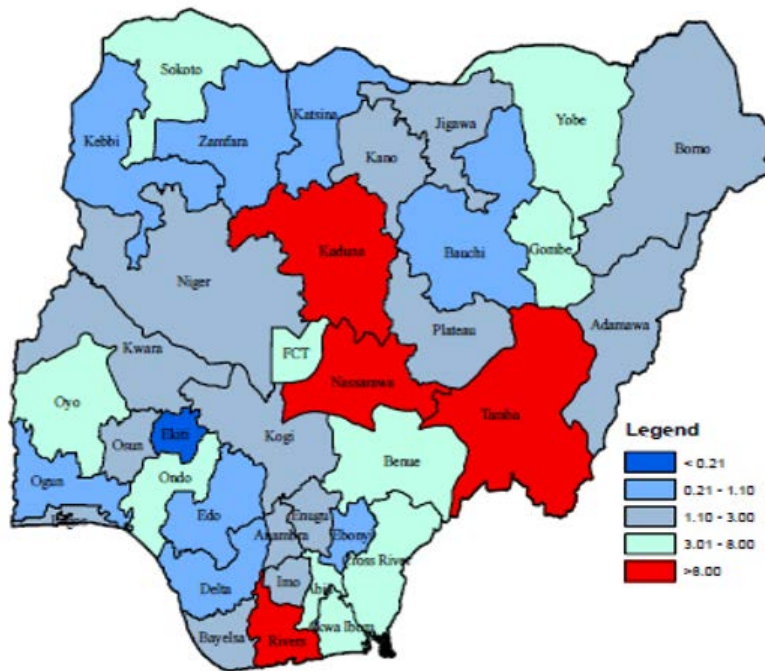


Fig. 2: Geographic Distribution of HIV Prevalence showing Ekiti having the least rate (NARHS 2012)

## 2 Materials and Methods

### 2.1 Data Collection

The data used for this study was collected from Ekiti State University Teaching Hospital. The data contains the ages, gender, local government area (LGA) and the HIV status of 200 patients reported at the hospital.

### 3 Method

The main objective of this study is to verify whether certain demographic characteristics really determine having or not having HIV among the reported cases in Ekiti state of Nigeria. The variable of interest for this study is whether HIV status which is either positive or negative. The logistic regression has been proven to be efficient in the analysis of data where response variable is categorical (Cox, 1972). In analyzing the data for studies, the outcome variable defined as "HIV status" is a categorical variable and thus requires use of the logistic regression approach. Logistic regression analysis extends the techniques of multiple regression analysis to a research situation in which the outcome variable is categorical. The model for logistic regression analysis assumes that the outcome variable,  $Y$  is categorical (e.g. dichotomous) but logistic regression analysis does not model this outcome variable directly. Rather, logistic regression analysis is based on probabilities associated with the values of  $Y$  (Fahrmeir and Tutz (2001), Hosmer and Lemeshow (2001)). Let assume  $Y$  is dichotomous, taking on values of 1 (i.e., the positive outcome, tested for HIV) and 0 (i.e., the negative outcome). In theory, the hypothetical population proportion of cases for which  $Y=1$  is defined as  $\pi$  i.e.  $P(Y=1) = \pi$ .

Logistic regression analysis is identical to the multiple regression analysis model except that the log-odd in favour of  $Y=1$  replaces the expected value of  $Y$ . The fundamental model underlying multiple regression analysis posits that a continuous outcome variable, a linear combination of a set of predictors and error (Ibrahim, Ipadeola, Adebayo, and Fatusi, 2013).

### 4 Logistic Regression Model

To fit a binary logistic regression model, a set of regression coefficients that predict the probability of the outcome of interest are estimated. The same logistic model can be written in different ways. Many texts in categorical statistics cover it (Agresti 1998), in addition to texts on logistic regression (Hosmer and Lemeshow 1989). Some analysts use the method with a different distribution function, the normal. In that case, it is called *probit analysis*. The version that shows what function of the probabilities results in a linear combination of parameters is:

$$\ln\left(\frac{\text{prob}(\text{event})}{1 - \text{prob}(\text{event})}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

The logit is the log of the odds that an event occurs. (The odds that an event occurs is the ratio of the number of people who experience the event to the number of people who do not. This is obtained when the probability that the event occurs is divided by the probability that the event does not occur.

The general linear logistic regression model is defined as:

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \text{logit}(\pi_i) = \beta_{i0} + \beta_{i1}x_{i1} + \dots + \beta_{ip}x_{ip} \quad (2)$$

where  $x_{i1}, x_{i2}, \dots, x_{ip}$  are continuous measurements corresponding covariates and/or dummy variables corresponding to factor levels and  $\beta_{i1}, \beta_{i2}, \dots, \beta_{ip}$  are the parameters.

$$\pi_i = \frac{e^{(\beta_{i0} + \beta_{i1}x_{i1} + \dots + \beta_{ip}x_{ip})}}{1 + e^{(\beta_{i0} + \beta_{i1}x_{i1} + \dots + \beta_{ip}x_{ip})}} = \frac{1}{1 + e^{-(\beta_{i0} + \beta_{i1}x_{i1} + \dots + \beta_{ip}x_{ip})}} \quad (3)$$

- 5  $\pi_i$  is the probability the  $i^{\text{th}}$  case experiences the event of interest.  
 $x_{ij}$  is the  $j^{\text{th}}$  predictor for the  $i^{\text{th}}$  case  
 $p$  is the number of predictors

There are two basic reasons underlying the development of logistic regression model. First, probabilities and odds obey multiplicative, rather than additive rules. However, taking the logarithm of the odds allows for the simpler additive model since logarithm converts multiplication into addition. And second, there is a (relatively) simple exponential transformation for converting log-odds back to probability.

## 6 Results and Discussion

All analyses were carried out using JMP 13 and STATA 14. Results are presented in the form of tables and charts.

**Table 1: Frequency distribution of HIV status of reported cases.**  
 Inserted in the parenthesis are the percentages

Variable	Negative	Positive	Patients
<b>Age</b>			
10-14	4(80)	1(20)	5
15-19	10(45.5)	12(54.5)	22
20-24	15(44.1)	19(55.9)	34
25-29	13(52)	12(48)	25
30-34	17(58.6)	12(41.4)	29
35-39	4(26.7)	11(73.3)	15
40-44	13(68.4)	6(31.6)	19
45-49	6(31.6)	13(68.4)	19
50-54	2(33.3)	4(66.7)	6
55-59	10(76.9)	3(23.1)	13
60-64	2(50)	2(50)	4
65-69	3(50)	3(50)	6
70-74	2(66.7)	1(33.3)	3
<b>Gender</b>			
Female	57(49.6)	58(50.4)	115
Male	44(51.8)	41(48.2)	85
<b>LGA</b>			
Ado-Ekiti	60(52.6)	54(47.4)	114
Ekiti-East	10(55.6)	8(44.4)	18
Ijero	3(33.3)	6(66.7)	9
Emure	0(0.0)	3(100)	3
Ekiti-South	1(25)	3(75)	4
Ekiti-West	1(33.3)	2(66.7)	3
Efon	9(56.3)	7(43.8)	16
Ise-Orun	2(40)	3(60)	5
Gbonyin	4(57.1)	3(42.9)	7
Ilejemeje	3(60)	2(40)	5
Ikere	1(50)	1(50)	2

Ikole	1(25)	3(75)	4
Ido/Osi	3(100)	0(0)	3
Moba	2(66.7)	1(33.3)	3
Oye	0(0)	3(100)	3
Irepodun	1(100)	0(0)	1

It can be observed from table 1, that the people in age group 35-39 have the highest frequency of HIV with 73% of the people in the group tested positive. The percentage tested positive among females is higher (50%) than the males (48%). All the patients tested in Emure and Oye local government areas are positive.

**Table 2: Whole Model Test**

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	9.68963	17.00000	19.37925	0.30720
Full	128.92981			
Reduced	138.61944			

**Table 3: Lack of Fit test**

Source	DF	-LogLikelihood	ChiSquare	Prob>ChiSq
Lack Of Fit	126.00000	83.79247	167.58494	
Saturated	143.00000	45.13734		
Fitted	17.00000	128.92981	0.00781	

**Table 4: Parameter Estimates**

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.3646965	792.25058	0.00000	0.99963
AGE	0.00663594	0.0104364	0.40430	0.52488
GENDER[female]	-0.0114739	0.1511445	0.00576	0.93949
LGA[Ado-Ekiti]	0.24721856	792.25051	0.00000	0.99975
LGA[Efon Alaaye]	0.40065814	792.25063	0.00000	0.99960
LGA[Ekiti East]	0.35551463	792.25062	0.00000	0.99964
LGA[Ekiti South-West]	-0.9786394	792.25124	0.00000	0.99901
LGA[Ekiti west]	-0.5599429	792.25132	0.00000	0.99944
LGA[Emure]	-18.048941	4905.9107	0.00001	0.99706
LGA[Gbonyin]	0.39866584	792.25081	0.00000	0.99960
LGA[Ido/Osi]	18.3181767	4895.0305	0.00001	0.99701
LGA[Ijero]	-0.536052	792.25077	0.00000	0.99946
LGA[Ikere]	0.04465154	792.25163	0.00000	0.99996
LGA[Ikole]	-0.9315604	792.25123	0.00000	0.99906
LGA[Ilejemeje]	0.55598799	792.25095	0.00000	0.99944
LGA[Irepodun]	18.3600793	8426.3692	0.00000	0.99826
LGA[Ise Orun]	-0.3239047	792.25096	0.00000	0.99967
LGA[Moba]	0.81409492	792.25132	0.00000	0.99918

Table 5: Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AGE	1	1.00000	0.40582555	0.52410
GENDER	1	1.00000	0.0057631	0.93949
LGA	15	15.00000	18.9523154	0.21590

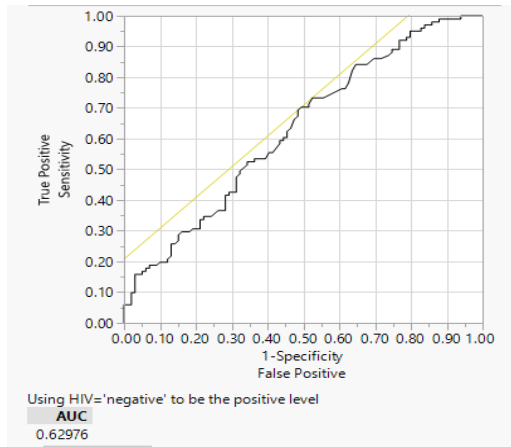


Fig. 3: Receiver Operating Characteristic

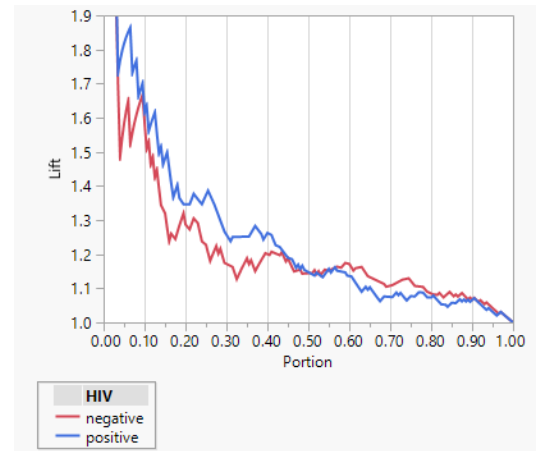


Fig. 4: Lift Curve

The Whole Model Test report shows if the model fits better than constant response probabilities. From the whole model test in table 2, the p-value is not significant, indicating that the factors age, gender and local government of the selected people have no significant effect on whether the people have HIV or not. Lack of fit is significant in table 3, while table 4 reveals the non-significant effect of all the categories of the predictors at 5% level of significant. It is discovered in the receiver operating curve (ROC) that the factors do not go beyond the threshold value, this is confirmed using DeLong theory that says: A test with no predictive ability produces a curve that follows the diagonal of the grid (DeLong, DeLong and Clarke-Pearson, 1988). This means that the factors do not contribute significantly to the people's HIV status. The predictive ability of the partition model as displayed by a lift curve in fig 4, shows that the considered factors cannot predict reliable probabilities for the response variable since it indicates a low lift curve.

## 7 Conclusion and Recommendation

It can be concluded from the results above that inasmuch that since none of the factors is significantly contributing to the HIV status of the selected people in Ekiti state, the fact about the state having the lowest rate in the number of people living with HIV in Nigeria is true and that it is the safest state to live in throughout the country taking HIV prevalence into consideration. It is therefore recommended that other researchers interested in the prevalence of any disease should not just dwell on the information received on media but establish it through data collection and appropriate inferential statistical analysis.

## REFERENCES

- [1] Agresti, A. (1990). *Categorical Data Analysis*, New York: John Wiley and Sons, Inc.

- [2] Cox, D. R. (1972). "Regression Models and Life Tables (with Discussion)." *Journal of the Royal Statistical Society, Series B* 34:187—220
- [3] David Hosmer and Stanley Lemeshow. (2001). *Applied Logistic Regression* (second edition), John Wiley & Sons: New York. ISBN: 0471208264 / 0-471-20826-4
- [4] DeLong, E., DeLong, D, and Clarke-Pearson, D.L. (1988). "Comparing the Areas Under Two or more Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach," *Biometrics* 44, 837–845.
- [5] Fahrmeir, L., Tutz, G., (2001). *Multivariate Statistical Modelling based on Generalized Linear Models (3rd edition)*, Springer: New York
- [6] Federal Ministry of Health (2012). *National HIV/AIDS and Reproductive Health Survey Plus (NARHS+)*, Abuja
- [7] Hosmer, D.W. and Lemeshow, S. (1989). *Applied Logistic Regression*, New York: John Wiley and Sons.
- [8] Muhammad Ibrahim, Oladipupo Ipadeola, Samson Adebayo, Adesegun Fatusi (2013). *Socio-demographic Determinants of HIV Counseling and Testing Uptake among Young People in Nigeria*. *International Journal of Prevention and Treatment*. 2(3): 23-31
- [9] NACA (2017). *National Strategic Framework on HIV and AIDS: 2017-2021*. Retrieved Feb. 15, 2019, from <https://www.avert.org/professionals/hiv-around-world/sub-saharan-africa/nigeria>
- [10] National Population Commission (NPC) [Nigeria] and ICF International. (2014). *Nigeria Demographic and Health Survey 2013*. Abuja, Nigeria, and Rockville, Maryland, USA: NPC and ICF International.
- [11] UNAIDS (2017). *Ending the AIDS epidemic by 2030*. Retrieved Feb. 21, 2019, from <http://www.unaids.org>
- [12] UNAIDS (2018). *Trend of New HIV Infections*. Retrieved Feb. 21, 2019, from <http://aidsinfo.unaids.org>