

USE OF LOGISTIC REGRESSION IN ASSESSING RISK FACTORS OF HIV/AIDS IN NORTH CENTRAL ZONE OF NIGERIA

Odeyemi, Joseph Bamidele

Department of Mathematics and Statistics, Federal Polytechnic, Offa, Kwara State, Nigeria

ABSTRACT

This article assessed risk factors of HIV reported in the National HIV/AIDS and Reproductive Health Survey (NARHS) 2007. It was planned to determine risk factors that could be used to predict and control the prevalence of HIV/AIDS in North Central zone of Nigeria. This was done through the method of logistic regression. To this end, a subset of national data from the Federal Ministry of Health (FMOH) on HIV and reproductive health were used. The risk factors assessed include location, education, religion, marital status, gender and age which were eventually broken down into twenty one variables. The appropriateness of the logistic regression models obtained was also assessed by pseudo R-squared. The variables adjudged significant in the final model included Age 30-39, 40-49; Qu'ranic, Primary, Secondary, and Higher education; Islam religion; currently married; and rural dweller. Of these, education was found to be dominant factor. It was then concluded that Education, Age, Marital Status and Location but not Gender and Religion be used to model the control of HIV prevalence in the North Central zone of Nigeria.

KEYWORDS: Human Immunodeficiency Virus (HIV), Acquired Immunodeficiency Syndrome (AIDS), R- squared, Odds ratio, Logistic Regression

Received for Publication: 30/03/14

Accepted for Publication: 04/05/14

Corresponding Author: bamidele12odeyemi@gmail.com

INTRODUCTION

It has been established that there are three major routes of transmitting HIV in Nigeria: sexual contact, exposure to infected body fluids and from mother to child during pregnancy, delivery, or breastfeeding. Also reported was that most of HIV infections are acquired through unprotected sexual relations where one partner has. Commercial sex impact risk and sexual assault have been believed to have risk of HIV transmission. Body fluid is also reported to be the second most frequent mode of HIV transmission and this is through the blood and blood products. Blood transfusions with infected blood result in transmission of infection. In developed countries the risk of acquiring HIV from a blood transfusion is said to be extremely low but higher in developing countries. The third most common way identified is HIV transmission from mother to child during pregnancy, during delivery and, after delivery, through breastfeeding.

According to the Federal Ministry of Health and NARHS division, North Central Zone of Nigeria comprises of Kwara, Niger, Plateau, Benue, Kaduna, Nassarawa, Kogi states and Federal Capital Territory. The people living here are predominantly farmers. The vegetation of the area is savanna. The division of Nigeria land into two parts puts the zone as part of Northern Nigeria. Since this zone is located at the centre of the country, then one finds the Yoruba people of the south west geo-political zone in Kogi and Kwara states. The Fulani, Hausa, Nupe, Tiv, Edoma, Ebira and other smaller tribes are also found in this zone. It is then a multicultural and multilingual zone. The river Niger and Benue cut across this zone, the vegetation is savannah which permits fishing, cattle rearing and arable farming are part of the occupation found in the land. Majority of the people are rural dwellers with low standard of



All rights reserved

This work by [Wilolud Journals](http://www.wiloludjournals.com) is licensed under a [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/)

living and characterized by low income earners. Some of the states in the zone are categorized as educationally disadvantage state. It is also a multi-religious zone with Christianity, Islam, and Traditional religion as major religion.

Records had shown that sex is traditionally handled privately and in secrecy in Nigeria and discussion of sex with teenagers is often seen as inappropriate. Providing sex education for young people is therefore hindered by religion and cultural objections. It has then become difficult to educate students and pupils on preventive measures for HIV infection. Education as knowledge, enlightenment or wisdom is expected to transmit values and desirable attitudes from one person to another. Education will socialize individuals. Education being the process of teaching, training and learning especially in schools is to develop knowledge and skills. To be educated is to show evidence of knowledge and understanding. The North Central Zone of Nigeria is educationally backward compared with the South West Zone of Nigeria.

One of the results of previous surveys on HIV infection in Nigeria had put this zone and one of the states here as most HIV prevalence. This might be because of the low standard of living, with poor socio-economic amenities, low level of education and other risk factors. To address this problem and minimize the transmission of HIV infection the government of Nigeria took some decisive steps. One of the steps is a biennial survey which assesses the level/rate of transmission of this deadly virus in order to score itself on the workability, efficiency and effectiveness of its HIV/AIDS programme. The first of it kind being the NARHS 2007 survey which covered all parts of the country.

Statement of the Problem

Previous researchers have claimed that HIV has become an epidemic disease and a challenge to human health and socio- economic development in the world in general. Many people infected with HIV are not aware, hence, they walk about with the virus unknowingly (NARHS, 2012). Many people have not gone for test to know their HIV status. HIV/AIDS statistics have not shown any significant improvement. This can be seen in the following statistics of HIV prevalence rate; 2009 (3.5), 2010 (3.4) , 2011(3.3), 2012 (3.1). The worst of it is that people might not be sick at the initial stage of the infection until it has become AIDS and ready to kill. Besides, as at Thursday, July 19, 2012, in preparation for the July 22-27, 2012 International AIDS Conference in Washington DC, the Director-General of World Health Organization (WHO), Dr. Margaret Chan, said every year more than a million people in low and middle income countries start taking antiretroviral drugs, but for every person who starts treatment, another two are newly infected (The Guardian, 2012).

Objectives of the study

The general objective of this study is to assess the risk factors of HIV prevalence using logistic mode. Specifically, the objectives are:

- (i) To ascertain risk factors of HIV reported in NARHS, 2007
- (ii) To build logistic model for HIV prevalence
- (iii) To estimate Pseudo R-squared by the use of Pearson chi squared
- (iv) To determine the influence of each risk factor in logistic model

Literature Review

The report of NARHS, 2007 by the Federal Ministry of Health, Nigeria shows that life expectancy in Nigeria has remained low, and has declined in recent times, partially due to the effect of HIV and AIDS. The infant mortality (IMR) has remained high and estimated at 99 per 1000 live births while the under-five mortality rate (U5MR) is 191 per 1000 live births (UNICEF in FMOH, 2007).

HIV/AIDS has been described as a disease of the human immune system caused by the HIV. Immune Deficiency means a weakness in the body's system to fight diseases. The illness interferes with the immune system, making people with it much more likely to get infections. A blood test for HIV looks for HIV antibodies. If these antibodies



All rights reserved

This work by [Wilolud Journals](#) is licensed under a [Creative Commons Attribution 3.0 Unported License](#)

are found in the blood in the body, it means the body has HIV infection. People who have these antibodies are referred to as HIV-Positive or having HIV disease (www.wikipedia.org/wiki/hiv/aids and <http://www.who.int/gho>). In the Central Intelligence Agency (CIA) face book reported by Wikipedia a list of HIV prevalence rate of countries per tribes and people puts Nigeria in the second position worldwide with 3.7% and Cape Verde and Afghanistan 0.1% HIV prevalence rate respectively (<http://www.data.worldbank.org/Indicator>).

In the February 14, 2014 of This Day Live, it was claimed that in a new survey of NARHS, 2012, HIV/AIDS prevalence in Nigeria has dropped to 3.4% with Taraba and Rivers state on top of the list (www.thisdaylive.com). Each of the 36 state in Nigeria has HIV prevalence above 1% (www.nigeria.unfpa.org).

In the Punch of 3rd December, 2013 Dr. Grace Okotu claimed that female Policemen in Nigeria have higher HIV prevalence (www.punch.com). All these are indications that HIV/AIDS prevalence should be continually address using every possible means to reduce its incidence. One of the ways is by revalidating and disseminating every research carried out by government and non governmental organizations.

METHODOLOGY

Data for the study were a subset of national data for North Central Zone collected by the Federal Ministry of Health (FMOH). The factors examined include location, education, religion, marital status, sex and age. Test is the response whose outcome is either positive or negative. These characteristics and factors were of different levels. Location is of two levels (rural and urban), education is made up of five levels (none, qur'anic, primary, secondary and higher), religion was broken down to five levels (islam, protestant, catholic, traditional and others), age group was of six levels (15-19, 20-24, 25-29, 30-39, 40-49, and 50-64), marital status was divided into six levels (currently married, cohabiting, never married, separated, divorced and widowed) and sex was of two levels (male and female). The data were presented state by state. Statistical Package for Social Sciences (SPSS) was used to process the data.

In the original data Benue state was coded 1, FCT 2, Kogi State 3, Kwara State 4, Nasarawa State 5, Niger State 6, and Plateau State 7; Under location- rural area was coded 1 and urban area was 2; Under HIV test- Positive test was 2 and Negative test was 1; Under education- None education was 0, Qur'anic was 1, Primary 2, Secondary education 3 and Higher education 4; Under religion- Other religion was 0, Islam was 1, Catholic was 2, Protestant was 3, Traditional was 4; For Age, 15-19 was 1, 20-24 was 2, 25-29 was 3, 30-39 was 4, 40-49 was 5, 50-64 was 6; For Sex, male was 1 and Female was 2; Lastly Marital Status, Currently married was 1, Living with sexual partner was 2, Never married was 3, Separated was 4, Divorced was 5 and Widowed was 6.

All these r levels except the state turned to r-1 variables so as to make the design matrix to be of full rank and have a reference. So Age turned to 5 variables, Education turned to 4 variables and Religion turned to 4 variables, Marital Status turned to 5 variables, Sex was still 1 variable, Location was still 1 variable and Test was still 1 variable. So there were twenty one variables altogether. All the variables were recoded. Location was coded 0 for rural and 1 for urban; Test was coded 0 for negative test and 1 for positive test; Education was coded 1 for quranic education and 0 otherwise, coded 1 for primary education and 0 otherwise, coded 1 for secondary education and 0 otherwise, coded 1 for higher education and 0 otherwise; Religion was coded 1 for islam religion and 0 otherwise, coded 1 for catholic religion and 0 otherwise, coded 1 for protester religion and 0 otherwise, coded 1 for traditional religion and 0 otherwise; Age was coded 1 for age 20-24 and 0 otherwise, coded 1 for age 25-29 and 0 otherwise, coded 1 for age 30-39 and 0 otherwise, coded 1 for age 40-49 and 0 otherwise, coded 1 for age 50-64 and 0 otherwise. Applying the forward conditional method whereby variables entered into the model one after the other and assessed significant or not by SPSS. Those variables not significant were dropped by the software.

All rights reserved



This work by [Wilolud Journals](http://www.wiloludjournals.com) is licensed under a [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/)

Table 1: Variable and Code

Variable	Response	Reference
Test (HIV)	1 (negative)	0 (positive)
Location	1 (rural)	0 (urban)
Gender	1 (Female)	0 (male)
Qu'ranic education	1	0
Primary education	1	0
Secondary education	1	0
Higher education	1	0
Islam religion	1	0
Catholic religion	1	0
Protestant religion	1	0
Traditional religion	1	0
Other religion	1	0
Age 20-24	1	0
Age 25-29	1	0
Age 30-39	1	0
Age 40-49	1	0
Age 50-64	1	0
Currently married	1	0
Living with sexual partner	1	0
Separated	1	0
Widowed	1	0
Divorced	1	0

The reference point for education is no education, for religion it is no religion, for age it is 15-19 and for marital status it is never married.

Data Analysis

The results of the computer processing are presented in the following tables below:

Table 2 : Initial Equation Containing no Variable

Constant	B	S.E	Wald	Df	Sig.	Exp(B)
	0.228	0.044	26.380	1	0.000	1.257

Here no variable is involved in the model



Table 3: Initial Assessment of Variables not in the Equation

Variable	Score	Df	Sig.	Reference
Sex (female)	9.391	1	0.002*	Male
Location (rural)	80.624	1	0.000*	urban
Age (agegroup2)	0.187	1	0.665	agegroup1
Age (agegroup3)	0.955	1	0.328	agegroup1
Age (agegroup4)	8.408	1	0.004*	agegroup1
Age (agegroup5)	0.042	1	0.838	agegroup1
Age (agegroup6)	3.321	1	0.072	agegroup1
Education (qur'anic)	0.826	1	0.364	no education
Education (primary)	1.132	1	0.287	no education
Education (secondary)	57.967	1	0.000*	no education
Education (higher)	86.967	1	0.000*	no education
Religion (islam)	44.868	1	0.000*	no religion
Religion (protestant)	36.994	1	0.000*	no religion
Religion (catholic)	3.441	1	0.064	no religion
Religion (traditional)	0.072	1	0.788	no religion
Religion (other religion)	0.079	1	0.779	no religion
Marital Status (currently married)	2.259	1	0.133	never married
Marital Status (living with sexual partner)	0.173	1	0.678	never married
Marital Status (separated)	0.018	1	0.893	never married
Marital Status (divorced)	0.570	1	0.450	never married
Marital status (widowed)	5.468	1	0.019*	never married

These are the initial assessment of the variables to be used in the model. Those that are significant at the initial assessment are clearly shown.

Table 4: Final Variables in the Equation (Step 9)

Variable	B	S.E	WALD	DF	SIG.	EXP(B)
Location	0.461	0.114	16.240	1	0.000	1.586
Agegroup4	0.426	0.128	11.054	1	0.001	1.531
Agegroup5	0.348	0.148	5.526	1	0.019	1.417
Qu'ranic education	0.839	0.233	12.919	1	0.000	2.315
Primary education	0.858	0.138	38.442	1	0.000	2.358
Secondary education	1.490	0.140	113.194	1	0.000	4.438
Higher education	2.199	0.216	103.495	1	0.000	9.014
Islam religion	-0.235	0.103	5.249	1	0.022	0.790
Currently married	0.246	0.109	5.088	1	0.024	1.279
Constant	-1.005	0.150	44.708	1	000	0.366

Variable(s) entered on step 1: higher. Variable(s) entered on step 2: secondary Variable(s) entered on step 3: primary. Variable(s) entered on step 4: location Variable(s) entered on step 5: agegroup4 Variable(s) entered on step 6: quranic Variable(s) entered on step 7: agegroup5 Variable(s) entered on step 8: Islam Variable(s) entered on step 9: currently married.



At step 9 only nine of the 21 variables were significant for the model and other were not.

Computation of R-squared

$$\chi_0^2 = 908 \quad R^2 = \frac{\chi_R^2}{\chi_0^2} = \frac{\chi_0^2 - \chi_E^2}{\chi_0^2}$$

Table 5: Pseudo R squared (model assessment)

s/no	χ_r^2	R_i^2	Difference	Cum. freq.	Percentage
1	86.668	0.0954	-	0.0954	23.19
2	195.564	0.2154	0.12	0.2154	29.17
3	232.162	0.2557	0.04	0.2554	9.72
4	253.171	0.2788	0.023	0.2784	5.59
5	266.82	0.2939	0.015	0.2934	3.65
6	276.107	0.3041	0.01	0.3034	2.43
7	302.02	0.3326	0.029	0.3324	7.05
8	327.096	0.3602	0.028	0.3608	6.81
9	373.348	0.4112	0.051	0.4114	12.4

Pseudo R squared of 0.41 shows a good model for HIV prevalence using the nine factor levels. Education accounted for 0.21 of this. This is 51.56% of the entire variation accounted for by the model.

RESULTS AND DISCUSSION

The summary of the results could be seen in Tables 4 and 5. The final step of the processing showed the highest variability which the explanatory variables could explain. These show the significance of the identified factors. A study of the Pseudo R-squared shows gradual increase from one level to another level. The partial pseudo R-squared indicates the contribution of each factor level as used in fitting the model. It is also noticed in table 4 that only nine levels out of twenty one were significant statistically. This is to say every one of these nine levels influenced HIV/AIDS prevalence in the zone. Of these, four are in education, two in age, and one in religion, one in location, one in marital status but none in gender. This shows the importance of education and age in HIV/AIDS prevalence.

The odds ratio of successes shown in the last column of table 4 indicates the significance of each of the levels. These values represent the number of times non cases of HIV prevalence are higher than cases of HIV prevalence. These also provide indirect estimate of North Central populace relative risk to HIV prevalence. In this study an odds ratio less than 1 indicates reduced odds of not having HIV prevalence among people of North Central zone while a



value greater than 1 indicates increased odds of not having HIV prevalence in North Central Zone of Nigeria. Therefore, odds ratio of 0.79 for Islam shows that it is a special risk factor.

A Pseudo R squared of 41% in table 5 is showing a significant assessment of the explanation of the variability by the model. The influence of the factor levels are shown in percentages in the table.

CONCLUSION

From this study those between the age of 30 and 49 with formal education are crucial in the fight against HIV/AIDS in this zone. This is consistent with the findings of NARHS, 2007 that those with lesser education are HIV risk population. Therefore, government should take education of its citizens more seriously. In addition, it was discovered that gender is of no statistical importance in this study, hence, sex of individual does not matter as opposed to the results of NARHS, 2007 that it was more prevalence among females. Of note is that there was reduction in the odds of no infection in Islam religion which supported the report of NARHS, 2007 that HIV prevalence was more among Muslim women.

RECOMMENDATIONS

In view of the findings above, all of Location, Marital Status, Age, and Education influence HIV prevalence and should be used to model the fight against HIV. However, Gender should not be used to model the fight against it. As Islam was discovered to be a special risk factor then campaign against the disease should be in the worship centres of different religions.

REFERENCES

- Berkson, J. (1944). Application of the logistic function to bioassay, *Journal of the American Statistical Association*, 39, 357-365
- Berkson, J. (1955). Maximum likelihood and minimum chi squared estimates of the logistic Function. *Journal of the American Statistical Association*, 50, 130-162
- Berkson, J. (1980). Minimum Chi square, not maximum likelihood! *The Annals of Statistics* 8, 457-467.
- Federal Ministry of Health (2006). *National HIV/AIDS and Reproductive Health Survey 2005*, Federal Ministry of Health, Abuja
- Federal Ministry of Health (2007). *National HIV/AIDS and Reproductive Health Survey* Federal Ministry of Health, Abuja
- Lillian, C and John, A. (2012). How to check HIV transmission, by WHO, others. *The Guardian*, 29 (12,257), 80

All rights reserved



This work by [Wilolud Journals](#) is licensed under a [Creative Commons Attribution 3.0 Unported License](#)

Marlo, S. and Nancy, F, (2003). Dose Finding Using Biased Coin Up-and-Down Design and Isotonic Regression *Journal of the International Biometrics Society* .58, 171-177

Macello, P. and Kimberlee, G., (1993). *Principles of Biostatistics*, Published by Wadsworth. Publishing Company, Belmont, California, 427-437

National Action Committee on AIDS (2004). *HIV/AIDS National Strategic Framework* 2005-2009, Abuja, NACA.

UNAIDS/WHO (2010). AIDS Epidemic Update. Retrieved February 5, 2012, from <http://data.unaids.org/pub/epiupdate>

This DayLive (February 14,2014). From <http://thisdaylive.com/article>. Accessed February 14, 2014

Wayne W. D., (2005). *Biostatistics: A Foundation for Analysis in the Health Sciences*, 7th Edition. Published by John Wiley and Sons, Inc., New York, 315-318

WHO (2011). Global Report on Tuberculosis. Accessed April 20, 2012, from <http://www.who.int/tb/publications/global-report/2011>

WHO (2013). *Global Health Observatory*. Accessed April 24, 2014 from <http://www.who.int/gho>

WorldBank (2013). Data Repository Prevalence of HIV Total (% of population ages 15-49).

Accessed April 24, 2014 from <http://www.data.worldbank.org/indicator>

All rights reserved



This work by [Wilolud Journals](#) is licensed under a [Creative Commons Attribution 3.0 Unported License](#)