

# The Impact of Demographic and Risk Behavioural Factors on the Prevalence of HIV/AIDS in Butajira, Ethiopia

Kassaye Wudu Seid<sup>1,\*</sup>, Yoseph Tadesse Dessie

Department of Statistics, Arba Minch University, Arba Minch, Ethiopia

## Email address:

kassayewudu16@gmail.com (K. W. Seid)

\*Corresponding author

## To cite this article:

Kassaye Wudu Seid, Yoseph Tadesse Dessie. The Impact of Demographic and Risk Behavioural Factors on the Prevalence of HIV/AIDS in Butajira, Ethiopia. *American Journal of Theoretical and Applied Statistics*. Vol. 6, No. 2, 2017, pp. 100-107. doi: 10.11648/j.ajtas.20170602.15

**Received:** November 12, 2016; **Accepted:** February 27, 2017; **Published:** March 18, 2017

---

**Abstract:** The study aims at describing the impacts of some demographic and risk behavior factors. The data for this study were taken from Butajira Hospital and data were analyzed using SPSS. In this study, chi square test of association is employed to test the association between determinant factors and the prevalence of HIV/AIDS in Butajira, Ethiopia. The result indicates all the proposed variables have significant association with prevalence of HIV/AIDS. Finally a binary logistic regression model is adopted for the analysis of the impact of demographic and HIV related risk factors on client's status with HIV/AIDS. The result revealed that, client's age, marital status, education level, occupation, suspected exposure time and condom use found to be significant at 5% significance level, indicating strong effects on prevalence of HIV/AIDS. The probability that an adult in this age groups 24-31, 32-45 and more than 46 years are 1.26 times (odds=1.26), 2.864 times (odds=2.864) 3.945 times (odds=3.945) more likely than those individuals aged between 16 and 23. Educational levels of clients have a significant influence on HIV infection (p value <0.05), as the result individuals who enrolled in secondary education level are more affected than those other individuals (i.e. Odds =1.227). It was also observed that those clients who had no sexual practices in the past are less likely to be infected by HIV/AIDS than those clients had sexual practice before (odds 0.453).

**Keywords:** HIV/AIDS, Binary Logistic Regression Analysis, Butajira Hospital, Ethiopia

---

## 1. Introduction

HIV/AIDS was a global epidemic in the history of world. The HIV/AIDS epidemic has become a series health and development problem in many countries around the world. According to the joint United Nation program on HIV/AIDS (UNAIDS), an estimated 39.5 million people worldwide have living with HIV at the end of 2006, of which 20.8 million have found in Sub-Saharan Africa. When AIDS first started, on one could have predicted how the epidemic would starts to spread across the world and how many millions of lives it would change. There was no real idea what caused it and consequently no real idea how to protect against it [1].

We know from bitter experience that HIV was the cause of AIDS and that it can devastate families, communities and whole countries. We have seen the epidemic knock decades off countries' national development, widen the gulf between rich and poor nations and push already stigmatized groups

closer to the margins of society. We have living in an "international" society, and HIV has become the first truly "international" epidemic, easily crossing oceans and borders. In some cases this indicates, among other things, that people have beginning to change risky behavior patterns, because they have seen and known people who have been killed by AIDS. It can also indicate that a large number of people have dying of AIDS. Already, more than 30 million people around the world have died of AIDS-related in 2007 UNAIDS has estimated that over 33 million people have living with HIV/AIDS and 70% of this amount have concentrated in sub-Saharan Africa. At the end of 2009, there are 9 countries in Africa where more than one tenth of the adult population aged 15-49 is infected with HIV. In 2010, 2.7 million people have newly infected with HIV, and 1.8 million men, women and children died of AIDS-related causes. 34 million people around the world have now living with HIV [1].

The role of voluntary counseling and testing can play on HIV/AIDS prevention and control was to be multifaceted.

Voluntary counseling and testing (VCT) service plays role in the prevention of HIV/AIDS especially in the rounded endeavor to provide care and support for people living with HIV/AIDS. It was worthy to see some of the impacts of Demographic and HIV related risk behavior factors on the status of HIV/AIDS for the case of Butajira.

The impact of HIV/AIDS was felt first by the individual, their families and eventually it out wards towards firms, business and economics of the country. Many national & organization have providing home care and support, counseling and nutritional support to those people with the virus. The rapid spread of the HIV epidemic was predominantly due to poverty. Government of countries made efforts in supporting orphans providing voluntary counseling and testing services, schools based education programs, peer counseling services and ART the vast majority of PLWHA in suffer from this epidemic [2].

The first evidence of HIV infection was found in 1984, and the first ADIS case is reported in 1986. The overall prevalence of HIV in the population was 4.4%. The highest prevalence rate is found in the 15-49 age groups and particularly in urban setting, where the prevalence was high as 12.6%, as opposed to rural settings here prevalence was approximately 2.6% and rising. These rates indicate that approximately 1.5 million people have infected and living with HIV in Ethiopia in a population of 69 million. 54.5% of all people living with HIV have women and 96,000 have children under the age of fifteen. In 2003 alone, an estimated 114,690 Ethiopia died of ADIS related conditions and this increased the number of children who have lost one or both parents to ADIS to over half a million. In 2003, it was estimated ADIS will further reduce life expectancy at birth of Ethiopia by 4.6 years. It would also continue to increase infant mortality and, lower the population size and growth in Ethiopia. Ethiopia has estimated 2 million people living with HIV and the third highest number of infection in Africa, according to UNADIS. With a population of 83 million people and per 5 capital income of less than US \$100 annually, it was also one of the world's poorest countries. [www.etharc.org/oromia/resources/kit/ARTInfotoolkit.pdf](http://www.etharc.org/oromia/resources/kit/ARTInfotoolkit.pdf).

The prevalence of inconsistent condom use among people living with HIV/Aids was 28.7%, however evidenced greater adherence among men than among women with statistically significant difference and the daily use of alcohol was associated with inconsistent condom use and this intern facilitates the prevalence of HIV/AIDS [10].

## 2. Data and Methodology

### 2.1. Sources and Data

The data used in this study have been collected form secondary sources, which have obtained from Butajira hospital. The data used in this study come from the special records kept at (BRHP) the counseling and testing service center in Butajira. The study subjects are men and women clients who visited the center for VCT service between January to May 2006. A total of 102 people received the service during the specified period and

records of all these clients have used for this study. The data have analyzed using SPSS software Packages.

### 2.2. Sampling Design

#### *Sampling procedures*

The target populations for this study have people enrolled in year 2006 E.C in Butajira hospital. In this study the respondent based on their identification number was used. The researcher first randomly picks the first item or subject from the population. Then, the researcher will select each  $K^{\text{th}}$  subject from the list. The results have representative of the population unless certain characteristics of the population are repeated for every  $K^{\text{th}}$  individual, which was highly unlikely. The integer was typically selected so that the researcher obtains the correct sample size. For sample of size calculation that can form to different research situations. According to sample size determination formula

$$K = \frac{N}{n} = \frac{5636}{102} = 55.25 \sim 55$$

Then select a random number between one and 55 using the set of random table. We select the individual allocate number 11 and then go on to select every 55<sup>th</sup> person, i.e. 66, 121, 176, then 231, and so on. Where  $z=1.96$  is the upper  $\alpha/2$  points standard normal distribution with  $\alpha=0.05$  significance level. Accordingly, the sample size (n), with population size  $N=5636$  for the current study was 102.

### 2.3. Variables Considered Under Study

#### *The Dependent Variables*

The response /dependent variable in this study is prevalence of HIV/AIDS which dichotomized as 1 if he /she are positive and as 0 if he /she are negative.

$$Y_i = \begin{cases} 1 & \text{if an individual is +ve} \\ 0 & \text{if an individual is -ve} \end{cases} \quad (1)$$

#### *The Independent Variables*

The independent variables that are used in this study are classified as demographic variables and HIV related risk behavior variables.

- I. Demographic Variables
  - a. age
  - b. Sex
  - c. Marital status
  - d. Educational level
  - e. Occupation
- II. HIV related risk behavioral variables
  - a. Previously tested
  - b. Ever has sex with partner
  - c. Suspected exposure time
  - d. Condom use last 3 months
  - e. History of STI

### 2.4. Statistical Methods

#### 2.4.1. Chi-square Test of Independency

Chi-square test of independency used to test the

relationship between two categories exists or not exists. The objective of chi-square test of independency is to test whether there is association between two categorical variables the null-hypothesis to be tested is that there is no relationship between two variables.

$$\chi^2 = \sum \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (2)$$

Where:  $-E_{ij} \rightarrow$  the expected frequency.

$O_{ij} \rightarrow$  observed frequency.

*The hypothesis*

$H_0$ : there is no significance association among demographic and behavioral variables and prevalence of HIV/AIDS.

$H_1$ : there is no significance association among demographic and behavioral variables and prevalence of HIV/AIDS

Reject the null hypothesis ( $H_0$ ) if  $X^2_{cal}$  greater than  $X^2_{tab}$  or if p-value less than level of significance Otherwise we accept  $H_0$

*Assumption of chi-square*

- The sample must be randomly selected from the population.
- The population must be normally distributed for the variable under study.
- The observation must be independent of each other.

#### 2.4.2. The Logistic Regression Model

The basic aim of modeling is to derive a mathematical representation of the relationship between an observed response variable and a number of explanatory variables, together with a measure of the inherent uncertainty of any such relationship. Statistical models are essentially descriptive and, in as much as they are based on experimental or observational data, may be described as empirical data [3].

The dependent variable in logistic regression is usually dichotomous, that is, the dependent variable can take the value 1 with a probability of success  $\theta$ , or the value 0 with a probability of failure  $1 - \theta$ . This type of variable is called a Bernoulli (or binary) variable. The independent or predictor variables in logistic regression could be discrete, continuous or a mix of both. Logistic regression makes no assumption about the distribution of the independent variables. They do not have to be normally distributed, linearly related or have equal variance within each group [9]. In logistic regression the relationship between the predictor and response variable is not linear. Instead, the logistic regression function is used, which is the logit transformation of  $\theta$ . In logistic regression, a single outcome variable  $Y_i$  ( $i=1 \dots n$ ) follows a Bernoulli probability function that takes on the value 1 with probability  $\theta_i$  and 0 with probability  $1 - \theta_i$ . Then  $\theta_i$  varies over the observations as an inverse logistic function of a vector  $X_i$ , which includes a constant and  $k$  explanatory variables, that is,

$$\theta(x) = \frac{e^{(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})}}{1 + e^{(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_k x_{ki})}} \quad (3)$$

Where  $\beta_0$  is the constant of the equation and,  $\beta_i$ 's are the coefficients of the predictor variables.

The goal of logistic regression is to correctly predict the category of outcomes for individual cases using the most parsimonious model. Variables can be entered into the model in the order specified by the researcher or logistic regression can test the fit of the model after each coefficient is added or deleted, called stepwise regression. Stepwise regression is used in the exploratory phase of research but is not recommended for theory testing [4]. Theory testing is the testing of a priori theories or hypothesis of the relationships between variables. Exploratory testing makes no a-priori assumptions regarding the relationships between the variables, thus the goal is to discover relationships. Backward stepwise regression appears to be the preferred method of exploratory analyses, where the analysis begins with a full or saturated model and variables are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed.

#### 2.4.3. Binary Logistic Regression Model

Binomial (or binary) logistic regression is a form of regression, which is used when the dependent variable is dichotomous, such as presence/absence or success/failure, HIV positive/HIV negative and the independent variables are of any type. Logistic regression can be used to predict a dependent variable on the basis of continuous and/or categorical independent.

variables and to determine the percent of variance in the dependent variable explained by the independent variables; to rank the relative importance of the independent variables; to assess interaction effects; and to understand the impact of the covariate control variables. The binary logistic model was also used to analyze factor that affect the status of HIV/AIDS.

The dependent variable in this case is binary, which take the value of 1 for positive and 0 for negative. Binary logistic regression analysis the odd of success defined as to be the ratio of the probability of success to the probability of failure.

$$\text{Logit}(P_i) = \text{Log} \left\{ \frac{P_i}{1 - P_i} \right\} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} \quad (4)$$

Where:  $-\beta_0, \beta_1, \dots, \beta_k$  are the model parameters.

#### 2.4.4. Fitting the Logistic Regression Model Parameter Estimation

Since the logistic regression model is nonlinear an iterative algorithm is necessary for parameter estimation. Let  $p$  be the probability of success and it is equivalent to the probability that the response variable assumes value one.

$$P(Y = 1) = \frac{1}{1 + e^{-x'\beta}} \quad (5)$$

Then each observation (response) can be considered as an outcome of a Bernoulli trial. Hence for the  $i^{\text{th}}$  observation  $Y_i$  the Bernoulli distribution is

$$P(Y = y_i) = p^{y_i} (1 - p)^{1 - y_i} \quad (6)$$

Since the observations are assumed to be independent the likelihood function is obtained as the product of the terms given in the above expression as follows.

$$LP^{y_i} (1 - \Pi p)^{1 - y_i} = \Pi \left( \frac{1}{1 + e^{-x'\beta}} \right)^{y_i} \left( \frac{e^{-x'\beta}}{1 + e^{-x'\beta}} \right)^{1 - y_i} \quad (7)$$

The principle of maximum likelihood states that we use as our estimate of the value which maximize the expression the log likelihood is defined as.

$$\ln L = \sum y_i \ln \left( \frac{1}{1 + e^{-x'\beta}} \right) + \sum (1 - y_i) \ln \left( \frac{e^{-x'\beta}}{1 + e^{-x'\beta}} \right) \quad (8)$$

Hence, through maximization of the above log-likelihood function we can theoretically estimate the parameter vector  $\beta$ . But the equation is nonlinear in  $\beta$ , and as a result the estimates do not have a closed form expression. Therefore,  $\beta$  will be obtained by maximizing using a numerical iterative method [5], [8].

#### 2.4.5. Assessment of Goodness Fit of Logistic Regression Model

The goodness of fit or calibration of a model measures how well the model describes the response variable. Assessing goodness of fit involves investigating how close values predicted by the model with that of observed values. Once a model had been developed through the various steps indicated in the above selection, we now would like to know how effective the model is in describing the outcome of the variable. We need to assess the goodness of fit of the model [5].

##### i. Hosmer-Lemeshow Test

The test divides subjects into deciles based on predicted probabilities, and then computes a Chi-square from observed and expected frequencies. Then a probability ( $p$ ) value is computed from the chi-square distribution with 8 degrees of freedom to test the fit of the logistic model. The Hosmer-Lemeshow test statistics is given by

$$C^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{v_k} \quad (9)$$

$H_0$ : there is no significance different between observed value and expected value

$H_1$ : there is difference between observed value and expected value

Where:  $O_k$  is the observed number of events in the  $k^{\text{th}}$  group,

$E_k$  is expected number of events in the  $k^{\text{th}}$  group, and  $V_k$  is a variance correction factor for the  $k^{\text{th}}$  group.

If the H-L goodness-of-fit P value is greater than 0.05, we will not reject the null hypothesis that there is no difference between observed and model-predicted values, implying that the model estimates fit the data at an acceptable level. That is, well fitting models show non significance on the H - L goodness-of-fit test, indicating that model-prediction is not significantly different from the observed values. This does not mean that the model necessarily explains much of the variance in the dependent, only that however much or little it does explain is significant. As the sample size gets large, the H-L statistic can find smaller and smaller differences between observed and model-predicted values to be significant [6], [7].

##### ii. The Wald Statistic test

The Wald test is a way of testing the significance of particular explanatory variables in a statistical model. In logistic regression we have a binary outcome variable and one or more explanatory variables. For each explanatory variable in the model there are associated parameters. If for a particular explanatory variable, or group of explanatory variables, the Wald test is significant, then we would conclude that the parameters associated with these variables are not zero, so that the variables should be included in the model. If the Wald test is not significant then these explanatory variables can be omitted from the model. Wald  $\chi^2$  statistics used to test the significance of individual coefficients in the model and are calculated as follows:

The hypothesis to be tested is:

$H_0$ :  $\beta_j = 0$  versus  $H_A$ :  $\beta_j \neq 0$  at  $\alpha$  level of significance.

The Wald test statistics,  $Z$ , for this hypothesis is

$$Z^2 = \left( \frac{\beta_j}{se(\beta_j)} \right)^2 \quad (10)$$

$\beta_j$  is the square of the estimated regression coefficient and  $se(\beta_j)$  is the standard error of  $\beta_j$ . However, several authors have identified problems with the use of Wald statistic. Menard warns that for large coefficients, the standard error is inflated, lowering the Wald statistic (chi-square) value. The likelihood ratio test is more reliable for small sample sizes than the Wald test [5].

## 3. Result and Discussion

The data are analyzed using the Statistical Package for Social Science (SPSS) Summary statistics are used to describe a set of observations, the Chi-square test was conducted to assess the association between variables and logistic regression analysis was run to assess the impact of predictors (explanatory) variable on the dependent variable.

### 3.1. Descriptive Statistics

From table 1 in the appendix, it shows that from a total number of 102 respondents 42.2% are males and the rest 57.8% are females. And 57.8% were in age group 16-23 years, 23.5% were in age group 24-31 years, 6.9% were in

age group 32-45 years and 11.8% were aged 46 years and above. From a total number of 102 respondents, 81 (79.41%) are HIV negative and the rest 21 (20.59%) of them are HIV positive.

### 3.2. Chi-Square Test Results

In order to investigate whether there is significant association between the demographic, risk behavioral variables and the prevalence of HIV/AIDS chi square test was employed. If the p-value of Chi-square test is less than  $\alpha=0.05$ , this indicates that there is an association between the dependent and independent variables.

#### Hypothesis

H<sub>0</sub>: There is no association between dependent variable and independent variable.

H<sub>1</sub>: There is association between dependent variable and independent variable.

The decision based on this test is that reject the null hypothesis if p-value less than level of significance ( $\alpha$ ). The p-value for all predictor variables as shown table 1 in the appendix is less than 0.05, this indicates that all demographic and behavioral factors are significantly associated with the prevalence of HIV/AIDS in Butajira, SNNPR.

### 3.3. Assessment of Model Adequacy

Table 1. Assessing model results for omnibus test of model coefficient.

|        |       | Chi-square | Df | Sig. |
|--------|-------|------------|----|------|
| Step 1 | Step  | 49.796     | 22 | .001 |
|        | Block | 49.796     | 22 | .001 |
|        | Model | 49.796     | 22 | .001 |

From table 1 above, the p value is 0.001 at 22 df which is less than 0.05 level of significance. It indicated that, considering demographic and risk behavioral variables together, the overall binary logistic regression model is significantly predicted the prevalence of HIV/AIDS in Butajira.

Table 2. Model Summary for binary logistic regression.

| Step | -2 Log likelihood | Cox & Snell R <sup>2</sup> | Nagelkerke R <sup>2</sup> |
|------|-------------------|----------------------------|---------------------------|
| 1    | 61.505            | .386                       | .582                      |

From the Table 2 above, Thecox & snell R square. 38.6% And Nagelkerke R square 58.2% indicates that the overall model is significant. This means all demographic and risk behavioral variables are useful to predict prevalence of HIV/AIDS in the study area.

Table 3. Hosmer and Lemeshaw test of goodness of fit.

| Step | Chi-square | Df | Sig. |
|------|------------|----|------|
| 1    | 7.096      | 8  | .526 |

This test indicates the fitness of predicted and actual value of prevalence of HIV/AIDS. The result presented in table 3 above shows the inferential goodness-of-fittest statistics (The Hosmerand Lemeshow test) that yield a  $\chi^2$  vale of 7.096 at 8 degree of freedom with p value 0.526 and it is insignificant at 5% level of significant, suggesting that the model was fit to

the data well. This means that the null hypothesis of a good model fit to data was tenable. In other words the Hosmer and Lemeshow goodness of-fit test statistic is not significant indicating that we do not have an evidence to reject the null hypothesis, suggesting that the model fitted the data well. So, binary logistic regression analysis can predict prevalence of HIV/AIDS very well.

### 3.4. Results of Logistic Regression Analysis

The results in table 5 in appendix revealed that, client's age, marital status, education level, occupation, suspected exposure time and condom use found to be significant at 5% significance level, indicating strong effects on prevalence of HIV/AIDS. However, sex has not significant impact on prevalence of HIV/AIDS. For the age predictor, the trend of odds increases with age of adults. The odds of adults with their age group 24-31 is 1.260 (CI: 2.520, 3.904) which implies that the probability that an adult in this age group infected by HIV/AIDS is 1.26 times more likely than those individuals in age group 16-23. The odds of adults with their age group 32-45 is 2.864 (CI: 2.176, 3.314) which implies that the probability that an individual in this age group infected by HIV/AIDS is 2.864 times more likely than those individuals in age group 16-23.

The odds of adults with their age group 46 and more is 3.945 (CI: 2.485, 3.558) which implies that the probability that an adult whose age 46 and more is infected by HIV/AIDS is 3.945 times more likely as compared to those individuals in age group 16-23. Marital status is the other determinant factor that shows the significant impact on the prevalence of HIV/ADIS. The odds of individuals who divorced with their partner is 1.23 (CI: 0.00, 0.022) which implies that individuals in this status is 1.23 times more likely to be infected by HIV/AIDS as compared to those individuals who are single and the odds of individuals who lost their partner is 0.026 (CI: 0.001, 0.895) which implies that individuals in this status is 0.026 times less likely to be infected by HIV/AIDS as compared to those single individuals.

The odds of individuals who enrolled in primary education level is 0.54 (CI: 0.50, 0.861) which implies that the probability that an individual in this category infected by HIV/AIDS is 0.50 times less likely than those individuals enrolled in junior education level. The odds of adults who enrolled in secondary education level is 1.227 (CI: 0.157, 2.606) which implies that the probability that an individual enroll in secondary education infected by HIV/AIDS is 1.227 times more likely than those individuals enrolled in junior education and the odds of adults who enrolled in tertiary education level is 0.020 (CI: 0.001, 0.476) this implies that the probability that an individual enrolled in tertiary education level infected by HIV/AIDS is 0.02 times less likely than those individuals enrolled in junior education.

The odds of individuals for those teachers is 0.358 (CI: 0.037, 3.496) which implies that the probability that an individual in this category infected by HIV/AIDS is 0.358 times less likely than those students. The odds of adults who has work is 0.031 (CI: 0.001, 0.813) which implies that the

probability that an individual who are workless infected by HIV/AIDS is 0.031 times less likely than those students. And the odds of adults who work in other profession is 0.157 (CI: 0.011, 2.286) this implies that the probability that an individual works in other profession infected by HIV/AIDS is 0.157 times less likely than those students.

The odds of individuals who failed to tested for HIV/AIDS is 2.695 (CI: 0.458, 1.874) which implies that the probability that an individual in this category infected by HIV/AIDS is 2.695 times more likely than those individuals who tested their blood previously for HIV/AIDS. And the odds of individuals who had no sex previously with other partners is 0.453 (CI: 0.075, 2.728) which indicates that the probability that an individual who had no sex previously with other partners is 0.453 times less likely than those individuals who had sex experience with other partners.

## 4. Conclusions

This study is an attempt to examine the impact of some

demographic and risk behavior characteristic for the prevalence of HIV/AIDS in Butajira. The socio-demographic profile of in clients shows that the probability of individuals to be infected by HIV/AIDS is dramatically increased with the age of individuals. Females are more likely to be HIV positive than males. This pattern of female vulnerability to HIV/AIDS infection in Butajira is resulted in many other studies. Educational levels of clients have a significant influence on HIV infection ( $p$  value  $<0.05$ ), as the result individuals who enrolled in secondary education level are more affected than those other individuals (i.e. Odds =1.227). It was also observed that those clients who had no sexual practices in the past are less likely to be infected by HIV/AIDS than those clients had sexual practice before (odds 0.453). This clearly shows that sexual intercourse is the most important factor that puts people at risk of contracting HIV/AIDS. Furthermore the findings showed that individuals who failed to tested their blood for HIV/AIDS in the past are more likely to be infected by HIV/AIDS than those individuals who had test practice in the past (odds 2.695).

## Appendices

*Table 1. Frequency distribution tables.*

| Variables                 | Category          | Frequency | Percent | HIV status |           | $\chi^2$ value | Df | P-Value |
|---------------------------|-------------------|-----------|---------|------------|-----------|----------------|----|---------|
|                           |                   |           |         | Positive % | Negative% |                |    |         |
| Age                       | 16-23             | 59        | 57.8    | 22.81      | 77.19     | 39.654         | 3  | 0.048   |
|                           | 24-31             | 24        | 23.5    | 20.83      | 79.17     |                |    |         |
|                           | 32-45             | 7         | 6.9     | 14.29      | 85.71     |                |    |         |
|                           | $\geq 46$         | 12        | 11.8    | 16.67      | 83.33     |                |    |         |
| Sex                       | Female            | 59        | 57.8    | 20.34      | 79.6      | 0.878          | 1  | 0.135   |
|                           | Male              | 43        | 42.2    | 20.93      | 79.07     |                |    |         |
| marital status            | Single            | 44        | 43.1    | 13.64      | 86.36     | 4.415          | 3  | 0.0122  |
|                           | Married           | 36        | 35.3    | 25         | 75        |                |    |         |
|                           | Divorced          | 16        | 15.7    | 25         | 75        |                |    |         |
|                           | widowed           | 6         | 5.9     | 33.33      | 66.67     |                |    |         |
| education level           | Junior            | 11        | 10.78   | 19.00      | 81.00     | 13.270         | 3  | 0.010   |
|                           | Primary           | 21        | 20.59   | 14.29      | 85.71     |                |    |         |
|                           | Secondary         | 33        | 32.36   | 18.18      | 81.82     |                |    |         |
|                           | Tertiary          | 37        | 36.27   | 16.22      | 83.78     |                |    |         |
| Occupation                | student           | 21        | 20.59   | 33.33      | 66.67     | 3.590          | 3  | 0.031   |
|                           | teacher           | 11        | 10.78   | 12.12      | 87.88     |                |    |         |
|                           | No work           | 17        | 16.67   | 17.65      | 82.35     |                |    |         |
|                           | others            | 53        | 51.96   | 13.21      | 86.89     |                |    |         |
| previously tested         | No                | 43        | 42.16   | 20.93      | 79.07     | 7.409          | 1  | 0.015   |
|                           | Yes               | 59        | 57.84   | 20.34      | 79.66     |                |    |         |
| ever has sex with partner | No                | 24        | 23.5    | 12.5       | 87.5      | 1.811          | 1  | 0.017   |
|                           | Yes               | 78        | 76.5    | 23.08      | 76.92     |                |    |         |
| suspected exposure time   | 1-3 weeks         | 28        | 27.5    | 28.57      | 71.43     | 8.811          | 3  | 0.032   |
|                           | 4-8 weeks         | 21        | 20.6    | 23.81      | 76.19     |                |    |         |
|                           | 9-12 weeks        | 21        | 20.6    | 19.05      | 80.95     |                |    |         |
|                           | Not exposed       | 32        | 31.3    | 15.63      | 83.37     |                |    |         |
| Condom use last 3 month   | Never             | 20        | 19.61   | 25         | 75        | 7.845          | 3  | 0.041   |
|                           | Always            | 32        | 31.37   | 12.5       | 87.5      |                |    |         |
|                           | Sometimes         | 33        | 32.35   | 27.27      | 72.73     |                |    |         |
|                           | does not remember | 17        | 16.67   | 17.65      | 82.35     |                |    |         |
| history of STI            | Yes               | 68        | 66.67   | 17.65      | 82.35     | 10.629         | 1  | 0.001   |
|                           | No                | 34        | 33.33   | 26.47      | 73.53     |                |    |         |
| HIV status                | Negative          | 81        | 79.41   |            |           |                |    |         |
|                           | Positive          | 21        | 20.59   |            |           |                |    |         |

**Table 2.** Case Processing Summary.

| Unweighted Cases (a) | N   | Percent |
|----------------------|-----|---------|
| Included in Analysis | 102 | 100.0   |
| Missing cases        | 0   | 0.0     |
| Total                | 102 | 100     |

If weight is in effect, see classification table for the total number of cases.

**Table 3.** Contingency Table for Hosmer and Lemeshow Test.

|        |    | HIVstatus = negative |          | HIVstatus = positive |          | Total |
|--------|----|----------------------|----------|----------------------|----------|-------|
|        |    | Observed             | Expected | Observed             | Expected |       |
| Step 1 | 1  | 10                   | 9.991    | 0                    | .009     | 10    |
|        | 2  | 10                   | 9.973    | 0                    | .027     | 10    |
|        | 3  | 9                    | 9.935    | 1                    | .065     | 10    |
|        | 4  | 10                   | 9.895    | 0                    | .105     | 10    |
|        | 5  | 10                   | 9.824    | 0                    | .176     | 10    |
|        | 6  | 9                    | 9.605    | 1                    | .395     | 10    |
|        | 7  | 10                   | 9.306    | 0                    | .694     | 10    |
|        | 8  | 10                   | 8.753    | 0                    | 1.247    | 10    |
|        | 9  | 8                    | 7.078    | 2                    | 2.922    | 10    |
|        | 10 | 2                    | 3.640    | 9                    | 7.360    | 11    |

**Table 4.** Multicollinearity.

| Model                     | Tolerance | VIF   |
|---------------------------|-----------|-------|
| Age                       | .651      | 1.536 |
| Sex                       | .911      | 1.098 |
| Marital Status            | .706      | 1.416 |
| Education Level           | .947      | 1.056 |
| Occupation                | .660      | 1.515 |
| Previously tested         | .735      | 1.361 |
| Ever has sex with partner | .605      | 1.653 |
| Suspected exposure time   | .779      | 1.284 |
| Condom use last 3 months  | .813      | 1.230 |
| History of STI            | .897      | 1.115 |

a Dependent Variable: HIV/AIDS status

**Table 5.** Result of Binary Logistic Regression Analysis.

|                              | $\beta$ | S.E   | Wald   | Df | Sig.  | Exp( $\beta$ ) | 95%CI for Exp( $\beta$ ) |       |
|------------------------------|---------|-------|--------|----|-------|----------------|--------------------------|-------|
|                              |         |       |        |    |       |                | Lower                    | Upper |
| Age (ref 16-23)              |         |       | 6.426  | 3  | 0.039 |                |                          |       |
| age (24-31)                  | 5.295   | 2.230 | 5.638  | 1  | 0.458 | 1.260          | 2.520                    | 3.904 |
| age (32-45)                  | 5.149   | 2.231 | 5.329  | 1  | 0.725 | 2.864          | 2.176                    | 3.314 |
| age ( $\geq 46$ )            | 4.399   | 1.831 | 6.037  | 1  | 0.023 | 3.945          | 2.485                    | 3.558 |
| sex: reference female        |         |       |        |    |       |                |                          |       |
| (male)                       | 0.684   | 0.900 | 0.576  | 1  | 0.045 | 0.798          | 0.339                    | 1.569 |
| Maritalstatus (ref: single)  |         |       | 13.336 | 3  | 0.000 |                |                          |       |
| Married                      | -8.707  | 2.644 | 10.847 | 1  | 0.163 | 0.000          | 0.000                    | 0.029 |
| Divorced                     | -8.764  | 2.527 | 12.022 | 1  | 0.050 | 1.230          | 0.000                    | 0.022 |
| Widowed                      | -3.648  | 1.805 | 4.086  | 1  | 0.998 | 0.026          | 0.001                    | 0.895 |
| Educationlevel (ref: Junior) |         |       | 7.025  | 3  | 0.001 |                |                          |       |
| Primary                      | -0.616  | 1.217 | 0.256  | 1  | 0.134 | 0.540          | 0.50                     | 0.861 |
| Secondary                    | 0.204   | 1.050 | 0.038  | 1  | 0.059 | 1.227          | 0.157                    | 2.606 |
| Tertiary                     | -3.892  | 1.607 | 5.865  | 1  | 0.051 | 0.020          | 0.001                    | 0.476 |
| Occupation (ref: Student)    |         |       | 4.612  | 3  | 0.003 |                |                          |       |
| Teacher                      | -1.027  | 1.162 | 0.780  | 1  | 0.998 | 0.358          | 0.037                    | 3.496 |
| No occupation                | -3.460  | 1.659 | 4.348  | 1  | 0.137 | 0.031          | 0.001                    | 0.813 |
| Others                       | -1.848  | 1.36  | 1.834  | 1  | 0.006 | 0.157          | 0.011                    | 2.286 |
| Previously tested (ref: yes) |         |       |        |    |       |                |                          |       |
| No                           | 0.991   | 0.905 | 1.201  | 1  | 0.052 | 2.695          | 0.458                    | 1.874 |
| ever had sex (ref: yes)      |         |       |        |    |       |                |                          |       |
| No                           | -0.792  | 0.916 | 0.748  | 1  | 0.087 | 0.453          | 0.075                    | 2.728 |
| Suspected exp (ref 1-3week)  |         |       | 7.385  | 3  | 0.021 |                |                          |       |
| 4-8week                      | -3.163  | 1.317 | 5.771  | 1  | 0.016 | 0.842          | 0.003                    | 0.559 |
| 9-12week                     | -2.397  | 1.172 | 4.186  | 1  | 0.041 | 1.691          | 0.009                    | 0.904 |

|                           | $\beta$ | S.E   | Wald  | Df | Sig.  | Exp( $\beta$ ) | 95%CI for Exp( $\beta$ ) |       |
|---------------------------|---------|-------|-------|----|-------|----------------|--------------------------|-------|
|                           |         |       |       |    |       |                | Lower                    | Upper |
| Don't exposed             | -3.323  | 1.408 | 5.570 | 1  | 0.018 | 0.036          | 0.002                    | 0.569 |
| Condomuse (ref: never)    |         |       | 5.882 | 3  | 0.018 |                |                          |       |
| Always                    | 3.153   | 1.440 | 4.793 | 1  | 0.029 | 3.412          | 1.392                    | 2.919 |
| Sometime                  | 1.837   | 1.346 | 1.862 | 1  | 0.172 | 6.276          | 0.449                    | 1.805 |
| don't remember            | 3.446   | 1.639 | 4.420 | 1  | 0.036 | 3.376          | 1.263                    | 2.393 |
| history of STI (ref: yes) |         |       |       |    |       |                |                          |       |
| No                        | -0.272  | 0.938 | 0.084 | 1  | 0.771 | 0.762          | 0.121                    | 4.785 |
| Constant                  | 2.506   | 2.202 | 1.295 | 1  | 0.000 | 12.254         |                          |       |

## References

- [1] J. G. Laah and E. Ayiwulu. Socio-Demographic Characteristics of Patients Diagnosed with HIV/AIDS in NasarawaEggon. *AJMS* 2010; 2 (3): 114-120.
- [2] Derek P et. al. UNAIDS Second Independent Evaluation: A Country Visit to Ethiopia, Summary Report. Geneva, Switzerland; 2008.
- [3] Collet D. Modeling Binary data. Second edition Chapman and Hall: London; 1991.
- [4] Scott W. Menard. Applied Logistic Regression Analysis. Second edition, Sage: USA; 1997.
- [5] Agresti, A. An Introduction to categorical data Analysis. Second edition, John Wiley and Sons: New York; 1996.
- [6] Hosmer, D and Lemeshow, S. Applied Logistic Regression. First edition, John Wiley and Sons: New York; 1989.
- [7] Hosmer D and Lemeshow S. Applied Logistic Regression. Second edition, John Wiley and Sons: New York; 2000.
- [8] Connel A. A. Logistic Regression Models for Ordinal Response Variables. Second edition, Sage: Thousand Oaks; 2006.
- [9] Long, J. S. (1997). Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage.
- [10] Renata Karina Reis et al. Factors associated with inconsistent condom use among people living with HIV/Aids 2016; 69 (1): 40-46.