International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Modeling Behavioral Factors Associated with HIV Infection among Women Aged 15-49 Years in Rwanda

Felix Hagenimana[1], Dr. Marcel Ndengo[2], Dr. Joseph K. Mung'atu[3]

[1]Student, Jomo Kenyatta University of Agriculture and Technology, Department of Statistics and Actuarial Science, Kenya

[2]Lecturer, University of Rwanda, College of Science and Technology, Kigali, Rwanda

[3]Lecturer, Jomo Kenyatta University of Agriculture and Technology, Department of Statistics and Actuarial Science, Kenya

Abstract: ***As part of the global push to end the AIDS epidemic for all by 2030, Rwanda is accelerating reductions of AIDS-related deaths and new HIV infections through strategies for reaching out and providing HIV services to key populations. HIV prevalence is higher among women (3.7%) compared to men (2.2%). We aimed to determine and model the key behavioral factors associated with HIV infection among women in Rwanda.We used secondary data from the 2010 Rwanda Demographic and Health Survey (RDHS). We matched data of HIV status and women aged 15-49 years. The potential covariates were defined through conceptual framework and tested using bivariate analysisto assess association between each covariate and outcome. Then, we used multivariate logistic regression to determine potential behavioral factors that are considered to have an impact on the prevalence of the epidemic. Of over 6900 women interviewed in 2010 RDHS, 3.7% were HIV positive and 96.3% were HIV negative. In multivariate logistic regression, HIV prevalence was higher in women who were between 35 to 49 years old (OR:2.090, 95%CI:1.230-3.549). Those who had experienced at least one STI symptom had a higher prevalence of HIV infection (OR:2.920, 95%CI:1.897- 4.493). We also found that women who were divorced/separated and widowed (OR:3.976, 95%CI:1.788-8.839), women who had first sex at 20-39years (OR:3.950, 95%CI:2.474-6.308) and women who had at least one sex partners (OR: 2.963, 95%CI: 1.149-7.639) were at higher risk for HIV infection. The use of condom during every sexual intercoursehad greatly reduced the risk of transmission of HIV infection(OR:0.145, 95%CI:0.098-0.214). In the context of HIV and AIDS prevention, limiting the number of sexual partners and encouraging protected sex are crucial to combating the epidemic. Successful prevention strategies should focus on HIV education, treatment of sexually transmitted infections, and proper and consistent condom use using an outreach approach.***

Keywords: HIV infection, AIDS, Multivariate Logistic Regression Model, Behavioral factors, Women, RDHS

1. Introduction

1.1 Background of the study

HIV infection has become the pandemic affecting every region of the world and is one of the major causes of morbidity and mortality. This deadly virus has affected the developing countries more severely as they are lagging behind in fight against various infectious diseases. With a high case fatality rate, significant impact on health and society, lack of definite curative treatment or vaccine, HIV/AIDS pandemic is one of the most frightening health problems of this century[1].AIDS is a disease which is caused by a virus called human immunodeficiency (HIV), that weakening the immunity of a person[2]. The HIV/AIDS epidemic is one of the largest obstacles to development in many countries and is destroying the lives and livelihoods of millions of people around the world[3]. HIV/AIDS has firstly recognized in 1980's, since then the epidemic is spread out rapidly all over the world. The trend of the number of people living with HIV AIDS is growing substantially from year to year and reached to its high level[4]. According to the US Global Health Policy report, the number of people living with HIV in the world is estimated to be between [31.6–35.2] millions, and between [2.4–2.9] millions people were newly infected with HIV per year and more than 2 million people have died due to this epidemic[1].HIV/AIDS killed 1.8 million of people where 1.3 million of people died live in Africa [5]. AIDS is now the fourth leading cause of death worldwide and the leading cause of death in Sub Saharan Africa [6].

According to [5]the prevalence rate in sub-Sahara Africa was 5% whereas 3% was for Rwanda. Adult national HIV prevalence rate in Rwanda refer to the 2005 Rwanda demographic health survey (RDHS), indicates that HIV prevalence nationwide is 3.1% among adult aged 15-49 years and women HIV/AIDS prevalence are higher than men 3.6% and 2.3%, respectively[7]. Therefore, prevalence among affected is different depends on sex and geographical areas of cohabitation. HIV data collected in 2005 RDHS showed HIV prevalence to be more than three times higher in urban than in rural areas (7.1% and 2.3%, respectively) with infection levels peaking in and around the capital Kigali. The prevalence of HIV/AIDS in Kigali city was 7.3% (8% in women and 5.2% in men); higher than the prevalence in other provinces and rural areas[8]. The wide gender difference in prevalence was a significant revelation that highlighted the vulnerability of women compared with men[9]. According to 2005 antenatal clinic survey showed that 4.1% of pregnant women were HIV infected, with prevalence highest in Kigali (13%), but on average about 5% in the urban areas and little over 2% in rural areas[10].

1.2 Statement of the problem

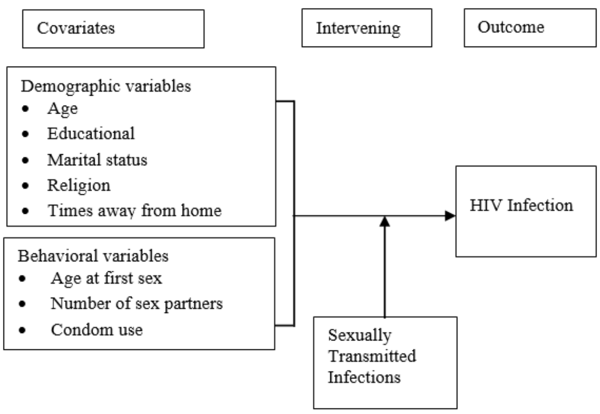
The predominant modes of HIV transmission identified are sexual intercourse, mother to child transmission, and transfusion with infected blood [6]. HIV infection is a major public health concern in Rwanda, where it is one of the major cause of mortality with negative socio-behavioral and economic consequences that affect national economy and human capital and even household wellbeing. HIV prevalence in Rwanda has remained stable (3%) since 2005 until 2010 in the people aged between 15 and 49 years old[11]. Recent Behavioral Surveillance Survey among Female Sex workers revealed a very high prevalence at 51% nationally, this key population is known to be an important mode of transmission of HIV and other STIs. This explains how HIV in Rwanda represents a mixed epidemic[10]. However, the behavioral factors associated with HIV infection among women in Rwanda were not yet well known and modelled. However, there is a need to model the key behavioral factors associated with HIV infection among women in Rwanda.

1.3 Objective of the Study

The objective of this study was to model the key behavioral factors associated with risk of HIV infection among women aged 15-49 years in Rwanda.

1.4 Conceptual framework of the study

The conceptual framework indicated a diagrammatic representation of the framework underlying this study: The HIV infection is associated with the different behavioral factors as shown in the following Figure.1



Source: Author

2. Methodology

2.1 Study settings

Rwanda is a small country and most populated in Africa with almost 11 million of population of whom 83% of them live in rural areas. Prevalence of HIV in adult is 3% and 2.2% in rural areas. This study is based on a secondary analysis of the data from the Rwanda demographic and Health Survey (RDHS 2010), which is a nationally and sub-nationally representative, conducted every five years to monitor demographic, socioeconomic, and health indicators. The DHS is conducted by the National Institute of Statistics of Rwanda in partnership with the Ministry of Health. This 2010 RDHS is a part of the worldwide Demographic and Health Survey (DHS) program, which is designed to provide data for monitoring the population and health situation in Rwanda. The 2010 RDHS is the fifth demographic and health survey conducted in Rwanda since 1992[11].

2.2 Study Design

The Rwanda Demographic and Health Survey(RDHS) is the first population-based, nationally representative survey in Rwanda to link individual HIV test results with the full set of behavioral, social, and demographic indicators included in the survey. In the RDHS 2010, a two stage stratified sampling design was used to sample 12,792 households from the country as a whole and for urban and rural areas in particular. Survey estimates are also reported for the provinces (South, West, North, and East) and for the City of Kigali. The sample was selected in two stages. In the first stage, 492 villages (also known as clusters or enumeration areas) were selected with probability proportional to the village size. The village size is the number of households residing in the village. For the second stage, the households were selected systematically from the lists of all survey's participators. All of the 492 clusters selected for the sample were surveyed for the RDHS 2010. The 2010 RDHS included HIV testing of over 6,900 women aged 15–49 and over 6,300 men aged 15–59 were identified as eligible for individual interviews and for HIV testing. Eligible women and men, HIV tests were conducted for 99 %of women and 98% of men.

Women who were interviewed in the subsample of households selected for the individual women's survey of the 2010 RDHS were asked to voluntarily provide blood for HIV testing. Therefore, the respondents' HIV test results can never be linked to identifying data. For women who were willing to be tested, drops of blood were drawn and dried on filter paper. Only an identification number (barcode) drawn at random was assigned to each specimen. Since no information containing personal identification accompanied the samples, it was not possible to inform the respondents of the result of their test. Analysis of the samples for HIV was carried out at the National Reference Laboratory (NRL).

The analysis presented in this study is based on sample of 6952 individual women that consented to the blood test result for HIV in Rwanda Demographic and Health Survey. As a result, blood test result data will be linked to survey data, and then the merged data file will be used to analyse the factors influencing HIV infection in Rwanda and unit of analysis here is women. The survey collected detailed information on fertility, family planning, childhood mortality, nutrition, maternal and child health, domestic violence, malaria, maternal mortality, awareness and behaviour regarding HIV/AIDS, HIV prevalence, malaria prevalence, and anemia. From these data we constructed a number of behavioural indicators that were likely to be associated with the risk of HIV infection[8].

Paper ID: SR201119235942 DOI: 10.21275/SR201119235942

2.3 Logistic Regression model

2.3.1 General Model: The logistic regression model for k covariates $(x_1, x_2, \ldots\ldots, x_k)$

Logisticregression is a mathematical modelling approach that can be used to describe the relationship of several independent variables to a dichotomousdependent variable[12]. Other modelling approaches are possible also, but logistic regression is by far the most popularmodelling procedure used to analyse dichotomous data. To explain the popularity of logistic regression, we show here the logistic function,which describes the mathematical form on which the logistic modelis based.

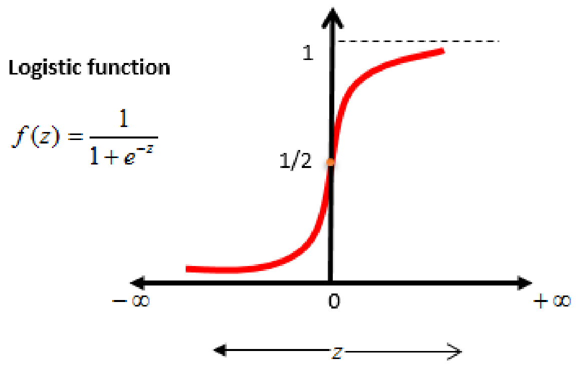


Figure 2: Logistic regression model shape

This function, called ***f*** (***z***), is given by 1 over 1 plus *e* to the minus *z*. We have plotted the values of this function as *z* varies from$-\infty$ *to* $+\infty$. Thus, as the graph describes, the rangeof *f*(*z*) is between 0 and 1, regardless of the value of ***z***. The fact that the logistic function *f* (*z*) ranges between 0 and 1 is the primary reason the logistic model is sopopular. The model is designed to describe a probability, which is always some number between 0 and 1. The logistic model, therefore, is set up to ensure thatwhatever estimate of risk we get, it will always be somenumber between 0 and 1. Thus, for the logistic model, we can never get a risk estimate either above 1 or below 0. This is not always true for other possible models, which is why the logistic model is often the first choice when a probability is to be estimated. To obtain the logistic model from the logistic function, we write *z* as the linear sum β_0 plus β_1 times X_1, plus $\beta_2$times X_1 , and so on to $\beta_k$times X_k, where the X's areindependent variables of interest and β_0 and the $\beta_i$are constant terms representing unknown parameters.

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k. \quad (1)$$

2.3.2 Probabilities, Odds, Odds ratios, and the Logit Transformation for Dichotomous Dependent Variables

For a dichotomous dependent variable, the numerical value of the variable is arbitrary, what is intrinsically interesting is whether the classification of cases into one or the other categories of the dependent variable can be predicted by the independent variable. Instead of trying to predict the arbitrary value associated with a category, it may be useful to predict the probability that a case will be classified into one as opposed to the other of the categories of the dependent variable.

Because the probability of being classified into the first or lower-valued category,$P(Y = 0)$, is equal to 1 minus the probability of being classified into the second or high-valued category,$P(Y = 1)$, it means$P(Y = 0) = 1 - P(Y = 1)$.

We could try to model the probability that $Y = 1$ as $P(Y = 1 = \alpha + \beta X$, but we could again run into the problem that although observed values of $P(Y = 1)$, must lie between 0 and 1, the predicted values may be less than 0 or greater than 1.

A step toward solving this problem would be to replace the probability that $Y = 1$ with the odds that$Y = 1$. The odds that$Y = 1$, written odds(Y=1), is the ratio of the probability that $Y = 1$ to the probability that $Y \neq 1$.The odds that $Y = 1$ is equal to $P(Y = 1)/[1 - P(Y = 1)]$. Unlike$P(Y = 1)$, the odds has no fixed maximum value, but like the probability, it has a minimum value of 0. One further transformation of the odds produces a variable that varies, in principle, from negative infinity to positive infinity. The natural logarithm of the odds,ln$\{\frac{P(Y=1)}{[1-P(Y=1)]}\}$, is called the *logit* of *Y*. The *logit* of *Y*, written$logit(Y)$, becomes negative and increasingly large in absolute value as the odds decrease from 1 to ward 0, and becomes increasingly large in the positive direction as the odds increase from 1 to infinity. If we use the logarithm of the odds that $Y = 1$ as our dependent variable, we no longer face the problem that the estimated probability may exceed the maximum or the minimum possible values for probability. The equation for the relationship between the dependent variable and independent variables then becomes $logit(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$[13]

We can convert $logit(Y)$ back to the odds by exponentiation,

Calculating$[odds\ that\ Y = 1] = e^{logit\ (Y)}$. This result in the equation

$$odds(Y = 1) = e^{\ln[odds\ (Y=1)]} = e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)} \quad (2)$$

We can then convert the odds back to the probability that $(Y = 1)$ by the formula

$P(Y = 1) = P(odds\ that\ Y = 1)/[1 + odds\ that\ (Y = 1]$.

This produces the equation

$$P(Y = 1) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}} \quad (3)$$

Let $\pi(X) = P(Y = 1)$ represent the probability of an event that depends on k covariates or independent variables. We have $\pi(X) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}$

$$logit(\pi(X)) = ln\left(\frac{\pi(X)}{1 - \pi(X)}\right)$$

$$= ln\left(\frac{\frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}}{1 - \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}}\right)$$

$$= ln\left(\frac{\frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}}{\frac{1}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}}}\right)$$

$$= ln\left(e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)}\right)$$

$$= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k \quad (4)$$

Paper ID: SR201119235942 DOI: 10.21275/SR201119235942

So, again, we see that the logit of the probability of an event given X is a simple linear function.

To summarize, the two basic equations of multivariate logistic regression are:

$$\pi(X) = \frac{e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_k X_k)}}{1+e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_k X_k)}}$$

Which gives the probabilities of outcome events given the covariate values $X_1, X_2, \ldots, X_p$

and $$logit(\pi(X)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

i.e. $ln\left(\frac{\pi(X)}{1-\pi(X)}\right) = \sum_{k=0}^{K} X_{ik}\beta_k \; i = 1, 2, \ldots, N$ (5)

Which shows that logistic regression is really just a standard linear regression model, once we transform the dichotomous outcome by the logit transform. This transform changes the range of $\pi(X)$ from 0 to 1 to $-\infty$ $to + \infty$, as usual for linear regression. Analogously to univariate logistic regression, the above equations are for mean probabilities, and each data point will have an error term. Once again, we assume that this error has mean zero, and that it follows a binomial distribution with mean $\pi(X)$ and variance$\pi(X)(1-\pi(X))$. Of course, now X is a vector, whereas before it was a scalar value.

It is important to understand that the probability, the odds, and the $logit$ are three different ways to express exactly the same thing. Of the three measures, the probability or the odds is probably the most easily understood? Mathematically, however, the $logit$ form of the probability best helps us to analyse dichotomous variables[14].

2.3.3 Interpretation of the β *coeffients* in Logistic Regression model

a) Interpretation of the intercept $\boldsymbol{\beta_0}$

Notice that regardless of the number of covariate values, if they are all set to zero, then we have:

$$\pi(X) = \frac{e^{\beta_0}}{1+e^{\beta_0}} \quad (6)$$

It is exactly the same as in the univariate case. So, the interpretation of β_0 remains the same as in the simpler case: β_0 sets the "baseline" event rate, through the above function, when all covariate values are set equal to zero.

b) Interpretation of the slopes $\boldsymbol{\beta_1, \beta_2, \ldots, \beta_k}$

Recall the effect on the probability of an event as X changes by one unit in the univariate case. There, we saw that the coefficient β_1 is such that e^{β_1} is the odds ratio for a unit change in X, and in general, for a change of Z units, the $OR(odd\ Ratio) = e^{Z\beta_1} = (e^{\beta_1})^Z$

Nothing much changes for the multivariate case, except:

1) When there is more than one independent variable, if all variables are completely uncorrelated with each other, then the interpretations of all coefficients are simple, and follow the above pattern:
   We have $OR = e^{Z\beta_i}$ for any variable X_i ,$i = 1,2,\ldots,k$ where the OR represents the Odds ratio for a change of size z for that variable.
2) When the variables are not uncorrelated, the interpretation is more difficult. It is common to say that $OR = e^{Z\beta_1}$ represents the odds ratio for a change of size z for that variable adjusted for the effects of the other variables. While this is essentially correct, we must keep in mind that confounding and co linearity can change and obscure these estimated relationships.

c) Estimating the β coefficients given data sets

The goal of logistic regression is to estimate the K +1 unknown parameters β in the equation(4). This is done with maximum likelihood estimation which entails finding the set of parameters for which the probability of the observed data is greatest. The maximum likelihood equation is derived from the probability distribution of the dependent variable. Since each y_i represents a binomial count in the i^{th} population, the joint probability density function of Y is:

$$f(y|\beta) = \prod_{i=1}^{N} \frac{n_i!}{y_i!(n_i-y_i)!}\pi_i^{y_i}(1-\pi_i)^{n_i-y_i} \quad (7)$$

For each population, there are $\binom{n_i}{y_i}$ different ways to arrange y_i successes from among n_i trials. Since the probability of a success for any one of the n_i trials is π_i , the probability of y_i successes $\pi_i^{y_i}$.

Likewise, the probability of $n_i - y_i$ failures is $(1-\pi i)ni-yi$

The joint probability density function in equation (7) expresses the values of y as a function of known, fixed the values for β. The likelihood function has the same form as the probability density function, except that the parameters of the function are reversed: the likelihood function expresses the values of β in terms of known, fixed the values for y.

Thus,

$$L(\beta|y) = \prod_{i=1}^{N} \frac{n_i!}{y_i!(n_i-y_i)!}\pi_i^{y_i}(1-\pi_i)^{n_i-y_i} \quad (8)$$

The maximum likelihood estimates are the values for β that maximize the likelihood function in Equation (8). The critical points of a function (maxima and minima) occur when the first derivative equals 0. If the second derivative evaluated at that point is less than zero, then the critical point is a maximum. Thus, finding the maximum likelihood estimates requires computing the first and second derivatives of the likelihood function. Attempting to take the derivative of Equation (8) with respect to β is a very difficult task due to the complexity of multiplicative terms. Fortunately, the likelihood equation can be considerably simplified. First, note that the factorial terms do not contain any of the π_i . As a result, they are essentially constants that can be ignored: maximizing the equation without the factorial terms will come to the same result as if they were included. Second, note that since $a^{x-y} = \frac{a^x}{a^y}$, and after rearranging terms, the equation to be maximized can be written as:

$$\prod_{i=1}^{N}\left(\frac{\pi_i}{1-\pi_i}\right)^{y_i}(1-\pi_i)^{n_i} \quad (9)$$

Note that taking exponential e to both sides of the equation (5)

$$\left(\frac{\pi_i}{1-\pi_i}\right) = e^{\sum_{k=0}^{K} X_{ik}\beta_k} \quad (10)$$

Which after solving for π_i becomes,

$$\pi_i = \left(\frac{e^{\sum_{k=0}^{K} X_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}}\right) \quad (11)$$

Substitution equation (10) for the first term and the equation (11) for the second term, equation (9) becomes:

$$\prod_{i=1}^{N}\left(e^{\sum_{k=0}^{K} X_{ik}\beta_k}\right)^{y_i}\left(1-\frac{e^{\sum_{k=0}^{K} X_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}}\right)^{n_i} \quad (12)$$

Simplifying the equation (11) can be written as:

$$\prod_{i=1}^{N} e^{y_i \sum_{k=0}^{K} X_{ik}\beta_k}\left(1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}\right)^{-n_i} \quad (13)$$

This is the kernel of the likelihood function to maximize. However, it is still cumbersome to differentiate and can be simplified a great deal further by taking its log. Since the logarithm is a monotonic function, any maximum of the likelihood function will also be a maximum of the log likelihood function and vice versa. Thus, taking the natural log of equation (13) yields the log likelihood function:

$$l(\beta)=\sum_{i=1}^{N} y_i\left(\sum_{k=0}^{K} X_{ik}\beta_k\right)-\ n_i\ ln\left(1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}\right) \quad (14)$$

To find the critical points of the log likelihood function, set the first derivative with respect to each β equal to zero. In differentiating equation (14) note that

$$\frac{\partial}{\partial \beta_k}\sum_{k=0}^{K} X_{ik}\beta_k = X_{ik} \quad (15)$$

Since the other terms in the summation do not depend on β_k and can thus be treated as constants. And then differentiating equation (14) with respect to each β_k we obtain:

$$\frac{\partial l(\beta)}{\partial \beta_k} = \sum_{i=1}^{N} y_i X_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}} \cdot \frac{\partial}{\partial \beta_k}\left(1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}\right)$$

$$= \sum_{i=1}^{N} y_i X_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}}\ e^{\sum_{k=0}^{K} X_{ik}\beta_k} \cdot \frac{\partial}{\partial \beta_k}\sum_{k=0}^{K} X_{ik}\beta_k$$

$$= \sum_{i=1}^{N} y_i X_{ik} - n_i \frac{1}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}}\ e^{\sum_{k=0}^{K} X_{ik}\beta_k}\ . X_{ik}$$

$$= \sum_{i=1}^{N} y_i X_{ik} - n_i\ \pi_i\ X_{ik} \quad (16)$$

The maximum likelihood estimates for β can be found by setting each of the K + 1 equations in equation (16) equal to zero and solving for each β_k Each such solution, if any exists, specifies a critical point either a maximum or a minimum. The critical point will be a maximum if the matrix of second partial derivatives is negative definite; that is, if every element on the diagonal of the matrix is less than zero. Another useful property of this matrix is that it forms the variance-covariance matrix of the parameter estimates. It is formed by differentiating each of the K + 1 equations in equation (15) a second time with respect to each element of β, denoted by $\beta_{k'}$. The general form of the matrix of second partial derivatives is

$$\frac{\partial^2 l(\beta)}{\partial \beta_k \partial \beta_{k'}} = \frac{\partial}{\partial \beta_{k'}}\sum_{i=1}^{N} y_i X_{ik} - n_i\ \pi_i\ X_{ik}$$

$$= \frac{\partial}{\partial \beta_{k'}}\sum_{i=1}^{N} -n_i\ \pi_i\ X_{ik}$$

$$= -\sum_{i=1}^{N} n_i\ \ X_{ik}\frac{\partial}{\partial \beta_{k'}}\left(\frac{e^{\sum_{k=0}^{K} X_{ik}\beta_k}}{1+e^{\sum_{k=0}^{K} X_{ik}\beta_k}}\right) \quad (17)$$

Then we use two differentiation rules.

First, a rule for differentiating exponential functions:

$$\frac{d}{dx}e^{u(x)} = e^{u(x)}\frac{d}{dx}u(x) \quad (18)$$

In our case, let $(x) = \sum_{k=0}^{K} X_{ik}\beta_k$. Second, the quotient rule for differentiating the quotient of two functions:

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-g'(x)f(x)}{[g(x)]^2} \quad (19)$$

$$\frac{d}{dx}\left(\frac{e^{u(x)}}{1+e^{u(x)}}\right) = \frac{(1+e^{u(x)}).e^{u(x)}\frac{d}{dx}u(x) - e^{u(x)}e^{u(x)}\frac{d}{dx}u(x)}{(1+e^{u(x)})^2} = \frac{e^{u(x)}\frac{d}{dx}u(x)}{(1+e^{u(x)})^2}$$

$$= \frac{e^{u(x)}}{1+e^{u(x)}} \cdot \frac{1}{1+e^{u(x)}}\frac{d}{dx}u(x) \quad (20)$$

Thus the equation (16) can be written as:

$$-\sum_{i=1}^{N} n_i\ \ X_{ik}\ \pi(X)(1-\pi(X))\ X_{ik'}. \ (21)$$

2.3.4 Confidence Interval for the Odds Ratio

The confidence interval for an Odds Ratio has the same general formula as the Confidence Interval for a population mean or population proportion.

Point Estimate $\pm$ Confidence Coefficient (CI)$*$Standard Error. The difference is that the confidence interval for the Odds Ratio is calculated on the natural log (ln) scale and then converted back to the original scale. The sampling distribution of the Odds Ratio(OR) is positively skewed; however, it is approximately normally distributed on the natural log scale. The lower and upper limits on the log scale = ln (OR) ± 1.96* SE ln (OR) = (LL, UL). Use the exponential function to find the CI limits on the original scale: EXP (LL), EXP (UL).

SE of ln (OR) = $\sqrt{\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)}$ then upper and lower limits on the ln scale will ln (OR) $\pm 1.96*\sqrt{\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)}$. Confidence interval of OR is obtained by using exponential function. Where a, b, c, d are respondents number in 2 X 2 table.

The coefficient β is the natural log of the odds ratio (odds of outcome among exposed versus odds of disease among unexposed), which has an approximate normal distribution. A 95% confidence interval for β is: (b – 1.96 SE (b), b + 1.96 SE (b). The population odds ratio, is: (exp [b – 1.96 SE (b)], exp [b + 1.96 SE (b)]). If this interval contains the value 1, then the relationship between outcome and covariate is not statistically significant at 5% level of significance.

2.3.5 Hypothesis testing for odds ratio

We tested the following hypotheses:

H0: β1 = β2=…… βn= 0 versus the alternative hypothesis

H1: βk ≠ 0 (at least one of the coefficient is different from 0)

The test statistic z = (b – 0)/SE (b) = b/SE(b) has a standard normal distribution if the null hypothesis is true.

2.4 Data analysis

We used secondary data from the 2010 Rwanda Demographic and Health Survey (RDHS). We matched HIV infection status and data of individual women aged 15-49 years old. The potential covariates were defined through conceptual framework and tested using bivariate analysis to assess association between each covariate and outcome. Non-collinearity (Pearson correlation r<0.5) covariates associated with HIV infection (p-value<0.10) were retained for multivariate model building. Then we used multivariate logistic regression to determine potential behavioural factors that are considered to have an impact on the prevalence of the epidemic. We ordered the remaining covariates from most-to-least import based on a priori knowledge, and used manual backward stepwise logistic regression retaining covariates that were associated with HIV infection (p-value<0.05). Stata V14.0 (College Station, Texas 77845 USA) software was used to perform data management and to generate the statistical analysis.We presented final model as odds ratio (OR) with 95% confidence intervals(CI).

3. Ethical statement

This study was a secondary analysis of the 2010 Rwanda Demographic Health Survey and as such, no ethical approval was required. We registered and requested for access to data from DHS on-line archive and received an approval to access and download de-identified DHS data files. All guidelines, including treating data as confidential and not making effort to identify individual respondents, were respected.

4. Results

Of over 6900 women interviewed in 2010 DHS, 3.7% were HIV positive and 96.3% were HIV negative. The majority of women (73.39%) were 15-34 years oldand 26.61% were 35-49 years old.The proportion of women interviewed (83%) had less than secondary education and 17% had secondary education or higher. Most of women (49.41%) were married and40.13% were single. Ofall women interviewed(96.86%) were Christians;22.3%have 2 or more timesaway from home in last 12 months;one in seven women (13.94%) had first sex at 6-13 years of ages, 0.35% had 2 or moresex partners (spouse excluded), 8.07% used condom during sexual intercourse and 2.23% of women had STIs in last 12 months (Table 1).

In the bivariate analysis in Table 3, age was associated with HIV infection, women who had 35-49 years oldwere more likely to have HIV infection than those who were between 15 and 24 years old (OR: 4.840, 95% CI: 3.441-6.804). Women who were divorced/separated and widowed had greater risk of HIV infection (OR: 7.396, 95%CI: 5.183 - 10.554). The bivariate analysis results also revealed that woman's times away from home in last 12 months was important predictor of HIV infection (OR: 1.154, 95%CI: 0.857-1.554). Woman's age at first sex was significantly associated with greater risk of HIV infection. The risk of HIV infection for women who had first sex atage of 6-19 years was seventeen times higher (OR:17.595, 95%CI: 9.049-34.211)than that women who had never had sex.Woman's number of sex partners(spouse excluded) was also important predictor of HIV infection (OR: 3.445, 95%CI: 2.417-4.909). STIs was associated with greater risk of HIV infection (OR: 8.975, 95%CI: 6.064-13.284). Finally, condom usewas associated with a decrease in the risk for HIV infection(OR:0.118, 95%CI: 0.084-0.165).

In the final multivariate logistic regressionin Table 3, all predictors were significant in bivariate analyses except woman's times away from home in last 12 months. Women who were between 35 to 49 years old were associated with higher HIV infection than those less than 34 years old (OR:2.090, 95%CI:1.230-3.549).Those who had experienced at least one STI symptom had a higher prevalence of HIV infection compared to those who had not (OR:2.920, 95%CI:1.897- 4.493).We also found that women who were married (OR:5.024, 95%CI:1.622-15.558), women who were divorced/separated (OR:3.976, 95%CI:1.788-8.839), women who were Muslim(OR:3.478, 95%CI:1.622-15.558), women who had first sex at 6-19 years (OR:2.920, 95%CI:1.897- 4.493), women who had first sex at 20-39years (OR:3.950, 95%CI:2.474-6.308) and women who had at least one sex partners (OR: 2.963, 95%CI: 1.149-7.639)were at higher risk for HIV infection.HIV infection was lower among women who reported using condom compared to women who did not (OR:0.145, 95%CI:0.098-0.214)

5. Discussion

Our findings demonstrated that HIV prevalence was 3.7% among woman in Rwanda, which is greater than the prevalence of men (2.2%) and the general population in Rwanda (3.0%)[8]. In contrast to the first two decades of the HIV pandemic, today women comprise about half of all adults living with HIV/AIDS globally[15]. Heterosexual transmission accounts for more than 80% of all new HIV infections in women. The majority of HIV-infected women live in sub-Saharan Africa where there is a substantial associated, concomitant epidemic of vertical transmission of HIV. Women acquire HIV infection at least 5-7 years earlier than men. HIV infection is 3-7 fold higher in adolescent women compared to adolescent boys in sub-Saharan Africa[16].

Among the women who had two or more partners in the past 12 months, almost two-thirds (63%) had overlapping (concurrent) sexual partnerships[8]. Concurrent sexual partnerships may increase the risk of HIV transmission because they allow the virus to pass quickly through multiple individuals.

The study found that HIV infection is particularly high among widows and those who are divorced or separated and have observed in 210 Demographic and health survey, the 16.6% of widows are HIV-positive[8].

STIs symptoms were associated with HIV infection. The relationship between STIs and HIV has been well established, most specifically with ulcerative STIs, which have been shown to be a frequent entry point of HIV virus[17]. HIV infection is also known to reduce the body’s immunity to fight against STIs[18][19].

The study revealed that the number of respondents reporting more than one partner in the past 12 months is very small (0.35%), therefore condom use by background characteristics is not noteworthy[8]. Numerous studies conducted over the past decade have demonstrated the steady increase in acceptability and use of male condoms by young people [20][21], particularly in settings where consistent messages are promoted and support for continued use is provided through access to free condoms[22][23]. However, significantly, very high levels of consistent condom use are required to reduce HIV incidence rates. A review of HIV risk studies [24] estimate that only about 20% of adolescents use male condoms consistently. While 70% of youth report having ever used a condom, about 50% report use of a condom at last coital encounter [22][25][26]. Among women, partnership type strongly influences condom use, with condoms generally viewed as less acceptable or desirable within long-term partnerships based on issues of love and trust [27][28], but acceptable in casual relationships. Various obstacles to condom use include negative beliefs about and attitudes toward condoms, often grounded in traditional gender constructions[29][30]. Some studies have found that young people may also associate condom use with promiscuity and sexually transmitted infections including HIV/AIDS[31][32]. Further, peer pressure or stigma about condom use inhibits actual use[33][34].

6. Conclusion

Given that most HIV infections are contracted through heterosexual contact, information on sexual behavior is important when designing and monitoring intervention programs to control the spread of the epidemic. In the context of HIV and AIDS prevention, limiting the number of sexual partners and encouraging protected sex are crucial to combating the epidemic.

Currently, a combination of strategies to ensure that key populations have access to a comprehensive package of HIV services as defined by the national program include:

- Provision of facility-based services package including systematic initiation of treatment as prevention, regular screening and testing for STIs and HIV, condom provision, provision of family planning services;
- Provision of community-based services such as HIV counselling and testing, STIs screening, condom distribution through outreach strategies;
- Linkage of community and health facility level interventions to ensure continuum of care;
- Organize support group of different categories of key population through peer education approach;
- Organize mass campaigns targeting key population groups to increasing their awareness and service utilization.

7. Acknowledgements

I extended my deepest appreciation to Dr. Joseph K. Mung'atu of Jomo Kenyatta University of Agriculture and Technology (JKUAT) and Dr. Marcel Ndengo of University of Rwanda (UR) who have been the ideal project supervisors. Their insightful criticisms and patient encouragements aided the writing of this project in innumerable ways.

References

[1] V. N. Baig, M. Swarnkar, A. K. Bhardwaj, M. Rathore, and A. Kasyap, "A study on socio-demographic profile and risk factors present in HIV infected patients attending art Centre in tertiary care hospital in Rajasthan, India.," *National Journal of Community Medicine*, pp. 3, 339, 2012.

[2] L. Johnson, "An introduction to the epidemiology of HIV/AIDS.," *Centre for Actuarial Research*, 2003.

[3] Ö. Katarina, "Socio-Economic Determinants of HIV in ZAMBIA.," *Nationalekonomiska Institutionen VID Lunds Universitet*, p. 187, 2009.

[4] T. Tsegaye, "Demographic, socio-economic and cultural determinants of HIV in Ethiopia.," *EKHR01*, 2010.

[5] UNAIDS, "Report on the Global AIDS Epidemic (Geneva: UNAIDS)," 2010.

[6] WHO, "The world health report 2002 - Reducing Risks, Promoting Healthy Life. Geneva: UNAIDS.," 2002.

[7] UNAIDS, "HIV/AIDS, Gender and young people. Kit publisher, Amsterdam," 2005.

[8] National Institute of Statistics of Rwanda (NISR) [Rwanda], Ministry of Health (MOH) [Rwanda], and ICF International, "Rwanda Demographic and Health Survey 2010," Calverton, MD: NISR, MOH, and ICF International, 2011.

[9] B. Williams, M. Lurie, A. Sturm, G. Garnett, D. Mkaya-Mwamburi, and S. Abdool Karim, "HIV discordance among migrant and non-migrant couples in South Africa.," *Paper presented at the XIII AIDS Conference. , Durban. Quoted in Lurie (2000)*, 2000.

[10] MoH, "Global AIDS Response Progress Report (GARPR).," 2014.

[11] NISR, MOH, and ICF International, "Rwanda Demographic and Health Survey2010: Final Report.," 2012.

[12] D. Kleinbaum and M. Klein, *Logistic Regression. Springer*. 2002.

[13] S. Menard, "Applied logistic regression analysis. A sage university paper.," 2002.

[14] R. Christensen, *Log-Linear Models and Logistic Regression. Springer*. 1997.

[15] UNAIDS. HIV and AIDS Estimates and Data,2001 - 2007., "Report on the Global AIDS Epidemic," cited; Available from: http://data.unaids.org/pub/GlobalReport/2008/JC1511_GR08_ExecutiveSummary_en.pdf, 2008.

[16] Abdool Karim Quarraisha, Sibeko Sengeziwe, and Baxter Cheryl, "Preventing HIV infection in Women – a Global Health Imperative!," *Published in final edited form as: Clin Infect Dis. 2010 May 15; 50(Suppl 3): S122–S129. doi:10.1086/651483.*

[17] A. Asiimwe, A. Koleros, and J. Chapman, "Understanding the dynamics of the HIV epidemic in Rwanda. Modeling the expected distribution of new HIV infections by exposure group. Kigali, Rwanda: National AIDS Control Commission of Rwanda and MEASURE Evaluation," p. 2009.

[18] Ministry of Health and Social Welfare of the United Republic of Tanzania, "[accessed 30 December 2016] Report of HIV Behavioral and Biological Surveillance Survey Among Female Sex Workers in Dar es Salaam,"

Available at: http://ihi.eprints.org/2793/1/NACP_Report_on_the_HIV_FSW.pdf.

[19] H. Vandenhoud, L. Langat, and J. Menten, "Prevalence of HIV and other sexually transmitted infections among female sex workers in Kisumu, Western Kenya, 1997 and 2008," *PLoS One. 2013; 8:e54953. [PubMed: 23372801].*

[20] J. Cleland and M. Ali, "Sexual abstinence, contraception, and condom use by young African women: a secondary analysis of survey data," *Lancet. Nov 18; 2006 368(9549):1788–93. [PubMed: 17113428].*

[21] DHS, "Demographic Health Survey website," www.measuredhs.com, 2005.

[22] A. Pettifor, H. Rees, and I. Kleinschmidt, "Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey," *AIDS (London, England). Sep 23; 2005 19(14):1525–34.*

[23] Goldstein S,, S. Usdin, E. Scheepers, and G. Japhet, "Communicating HIV and AIDS, what works? A report on the impact evaluation of Soul City's fourth series," *Journal of health communication. Jul- Aug; 2005 10(5):465–83. [PubMed: 16199389].*

[24] Eaton L,, Flisher AJ,, and Aaro LE, "Unsafe sexual behaviour in South African youth.," *Social science & medicine (1982),* vol. Jan; 2003 56(1):149–65.

[25] S. Hoffman, L. O'Sullivan, A. Harrison, C. Dolezal, and A. Monroe-Wise, "HIV risk behaviors and the context of sexual coercion in young adults' sexual interactions: results from a diary study in rural South Africa," *Sex Transm Dis. Jan; 2006 33(1):52–8. [PubMed: 16385222].*

[26] L. Simbayi, J. Chauveau, and O. Shisana, "Behavioural responses of South African youth to the HIV/ AIDS epidemic: a nationwide survey," *AIDS Care. Jul; 2004 16(5):605–18. [PubMed: 15223530].*

[27] S. Abdool Karim and Q. Abdool Karim, "The evolving HIV epidemic in South Africa," *Int J of Epidemiology*, pp. 31:37–40, 2002.

[28] D. Worth, "Sexual decision-making and AIDS: why condom promotion among vulnerable women is likely to fail. Studies in family planning," *Nov-Dec; 1989 20(6 Pt 1):297–307. [PubMed: 2623725].*

[29] A. Meyer-Weitz, P. Reddy, H. Van Den Borne, G. Kok, and J. Pietersen, "Determinants of Multi-Partner Behaviour of Male Patients with Sexually Transmitted Diseases in South Africa: Implications for Interventions," *International Journal of Men's Health. 2003; 2(2):149–62.*

[30] N. Morojele, J. Brook, and M. Kachieng'a, "Perceptions of sexual risk behaviours and substance abuse among adolescents in South Africa: a qualitative investigation," *AIDS Care. Apr; 2006 18(3): 215–9. [PubMed: 16546781].*

[31] G. Buga, D. Amoko, and D. Ncayiyana, "Sexual behaviour, contraceptive practice and reproductive health among school adolescents in rural Transkei," *South African medical journal = Suid- Afrikaanse tydskrif vir geneeskunde. May; 1996 86(5):523–7. [PubMed: 8711547].*

[32] C. Varga, "Sexual decision-making and negotiation in the midst of AIDS: youth in KwaZulu-Natal South Africa.," *Health Transition Review. 1997; 7(Supplement 3):45–67.*

[33] H. Boer and M. Mashamba, "Psychosocial correlates of HIV protection motivation among black adolescents in Venda, South Africa," *AIDS Educ Prev. Dec; 2005 17(6):590–602. [PubMed: 16398579].*

[34] C. MacPhail and C. Campbell, "I think condoms are good but, aai, I hate those things': condom use among adolescents and young people in a Southern African township," *Social science & medicine (1982). Jun; 2001 52(11):1613–27.*

Annexes:

Table 1: Description of demographic and behavioral characteristics of women interviewed in Rwanda, 2010 DHS

	Number (n)	Percentage (%)
HIV Status		
hiv negative	6,692	96.3
hiv positive	260	3.7
Woman's age		
15-24	2,943	42.33
25-34	2,159	31.06
35-49	1,850	26.61
Woman's education		
Secondary +	1,182	17.00
Less than secondary	5,770	83.00
Woman's marital status		
Never in union	2,790	40.13
Married/union	3,435	49.41
Divorced/separated/widowed	727	10.46
Woman's religion		
Christian	6,734	96.86
Muslim	90	1.29
Other	128	1.84
Woman's times away from home in last 12 months		
0	3,698	53.19
1	1,700	24.45
2+	1,554	22.35
Woman's age at first sex		
Never had sex	2,138	30.77
6-19	969	13.94
20-39	540	7.77
At first union	3,302	47.52
Woman's number of sex partners		
0	6,561	94.38
1	367	5.28
2+	24	0.35
Woman's condom use		
No	3,521	91.93
Yes	309	8.07
Woman had any STIS in last 12 months		
No	6,784	97.77
Yes	155	2.23

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Table 2: Collinearity screening

	w_age	w_edu	w_mar	w_reli~n	w_tima~y	w_ageF~x	w_numb~t	w_cond~e	w_STIs
w_age	1.0000								
w_edu	0.0089	1.0000							
w_mar	0.2058	0.0693	1.0000						
w_religion	-0.0060	-0.0484	0.0174	1.0000					
w_timaway	-0.0906	-0.1763	-0.0755	0.0406	1.0000				
w_ageFirst~x	0.2179	0.1485	0.2534	-0.0583	-0.1103	1.0000			
w_number_p~t	-0.1751	-0.0617	-0.3150	0.0594	0.1002	-0.3853	1.0000		
w_condom_use	-0.0663	-0.1139	-0.1350	0.0418	0.0611	-0.1584	0.2839	1.0000	
w_STIs	-0.0146	-0.0178	-0.0045	0.0623	0.0402	-0.0307	0.0606	0.0771	1.0000

Variable description: w_age: Woman's age, w-edu: Woman's education, w_mar: Woman's marital status, w_religion: Woman's religion, w_timaway: Woman's times away from home in last 12 months, w_ageFirst~x: Woman's age at first sex, w_number_p~t: Woman's number of sex partners(spouse excluded), w_condom_use: Woman's condom use, w_STIs: Woman had any STIS in last 12 months

Table 3: Bivariate and multivariate analyses of predictors of HIV among women in Rwanda, 2010 DHS

Predictors	Bivariate analysis[1]		Full multivariate logistic regression model[2]		Reduced/Final Model[3]	
	OR (95% CI)	p-value	OR (95% CI)	p-value	OR (95% CI)	p-value
Woman's age						
15-24	1.00		1.00		1.00	
25-34	2.676 (1.865-3.840)	0.000	1.183 (0.708-1.975)	0.521	1.180 (0.706- 1.973)	0.528
35-49	4.840 (3.441-6.804)	0.000	2.043 (1.203- 3.465)	0.008	2.090 (1.230-3.549)	0.006
Woman's education						
Secondary +	1.00		1.00		1.00	
Less than secondary	0.759 (0.562-1.027)	0.074	0.629 (0.409-0.966)	0.034	0.655 (0.430-0.998)	0.049
Woman's marital status						
Never in union	1.00		1.00		1.00	
Married/union	2.062 (1.484-2.865)	0.000	5.172 (1.671-16.010)	0.004	5.024 (1.622-15.558)	0.005
Divorced/separated/widowed	7.396 (5.183 -10.554)	0.000	4.256 (1.905-9.505)	0.000	3.976 (1.788-8.839)	0.001
Woman's religion						
Christian	1.00		1.00		1.00	
Muslim	4.092 (2.199-7.615)	0.000	3.405 (1.585-7.315)	0.002	3.478 (1.622-15.558)	0.001
Other	2.254 (1.167-4.353)	0.015	1.136 (0.438-2.944)	0.793	1.124 (1.788-8.839)	0.809
Woman's times away from home in last 12 months						
0	1.00		1.00		-	-
1	0.937 (0.688-1.275)	0.678	0.627 (0.404-0.973)	0.037	-	-
2+	1.154 (0.857-1.554)	0.347	0.938 (0.612-1.434)	0.764	-	-
Woman's age at first sex						
Never had sex	1.00		1.00		1.00	
6-19	17.595 (9.049-34.211)	0.000	2.993 (1.944-4.607)	0.000	2.920 (1.897- 4.493)	0.000
20-39	22.194 (11.188-44.0245)	0.000	3.986 (2.492-6.376)	0.000	3.950 (2.474-6.308)	0.000
At first union	8.791 (4.611-16.762)	0.000	1.00	-	1.00	
Woman's number of sex partners						
0	1.00		1.00		1.00	
1	3.445 (2.417-4.909)	0.000	2.940 (1.145-7.549)	0.025	2.963 (1.149-7.639)	0.025
2+	1.224 (0.165-9.106)	0.843	0.239 (0.023-2.493)	0.232	0.263 (0.026-2.708)	0.262
Woman's condom use						
No	1.00		1.00		1.00	
Yes	0.118(0.084-0.165)	0.000	0.142(0.096-0.210)	0.000	0.145(0.098-0.214)	0.000
Woman had any STIS in last 12 months						
No	1.00		1.00		1.00	
Yes	8.975 (6.064-13.284)	0.000	6.493 (3.928-10.732)	0.000	6.495 (3.930-10.735)	0.000

[1]In bivariate analysis, all odds ratios were adjusted for hiv status
[2]The full model included all variables with p-values<0.10 from bivariate analysis and Collinearity screening with Pearson correlation, |r|>0.5
[3]The final model included all variables with p-values<0.05 from the full model, and Manual backward stepwise logistic regression with p-value< 0.05) for reduced model