

**Report**

# A Bivariate Probit Model for Correlated Binary Data with Application to HIV and Male Circumcision

**Tabitha Wambui Njoroge, Samuel Musili Mwalili, Anthony Kibira Wanjoya**

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

**Email address:**

njorogewtabitha@gmail.com (T. W. Njoroge), samuel.mwalili@gmail.com (S. M. Mwalili), awanoya@gmail.com (A. K. Wanjoya)

**To cite this article:**Tabitha Wambui Njoroge, Samuel Musili Mwalili, Anthony Kibira Wanjoya. A Bivariate Probit Model for Correlated Binary Data with Application to HIV and Male Circumcision. *American Journal of Theoretical and Applied Statistics*. Vol. 4, No. 6, 2015, pp. 555-561. doi: 10.11648/j.ajtas.20150406.27

---

**Abstract:** As the HIV/AIDS epidemic continues to grow, it continues to be a huge threat to the social and economic well-being of a society. Studies show that the epidemic has significantly affected the development of Kenya. Numerous interventions by different bodies (e.g. the national government, international donors, civil society organizations) to prevent its spread continue to be put in place. Male Circumcision has been proven to reduce the risk of HIV transmission. A statistical model that shows the relationship between male circumcision and HIV prevalence is therefore of great importance as it can be used to bring out the inverse relationship between the two response variables and hence support male circumcision as an effective intervention for prevention of HIV spread. We use Bivariate Probit regression to model the correlation between Male Circumcision and HIV prevalence while looking into factors affecting both HIV and Male Circumcision.

**Keywords:** Bivariate Probit Model, Human Immunodeficiency Virus (HIV), Male Circumcision (MC), Correlation

---

## 1. Introduction

### 1.1. Background of the Study

According to UN Political Declaration on HIV/AIDS (2012), Kenya has the fourth largest HIV epidemic globally. The study showed that an estimated 1.6 million were infected with HIV and about 57,000 people died out of illnesses caused by AIDS. This has put pressure on the healthcare system as well as the economy of the country. By the mid 90's, HIV was one of the major causes of deaths in the country. There was a notable drop in the life expectancy in Kenya from 60 years in 1990 to 45.5 years in 2002, which was largely caused by AIDS [1].

Male circumcision is the surgical removal of the foreskin of the penis. The word 'circumcision comes from the Latin word 'circumcidere' which means 'to cut around' [2]. Ancient Egyptian paintings are evidence of its long existence. In Kenya, male circumcision is commonly practiced as a rite of passage into adulthood and also the promotion of hygiene. Different regions practice it at different levels, with Nyanza province practicing it the least. Several factors like cultural beliefs, religious beliefs, and social status can be attributed to

the discrepancy on the practice level in the different regions.

Male Circumcision has been proven to reduce the risk of HIV transmission. According to WHO, medical male circumcision reduces the risk of female-to-male sexual transmission of HIV by approximately 60%. The effect of having sufficient numbers of males circumcised could be as a herd immunity since preventing men from becoming infected will also protect their sex partners [3]. If 80% of men in the priority African countries were circumcised, 3.4 million HIV infections could be prevented over 15 years. [4]

Sufficient statistical models showing the correlation between Male circumcision and HIV prevalence do not exist. Several clinical trials have been performed that show evidence that heterosexual men who are circumcised are less likely to be infected by HIV compared to their uncircumcised counter-parts. Little effort has however been put to bring out this evidence in a mathematical model. In

Interventions to curb its spread is therefore encouraged worldwide, and any evidence of an effective intervention should be harnessed and a lot of effort be put into grounding the intervention. Three randomized controlled trials were conducted in Kenya, Uganda and South Africa which

produced findings that indicated male circumcision effectively reduces heterosexual HIV acquisition. The overwhelmingly 60% reduction of risk of HIV acquisition is regarded as effective as the long hoped for AIDS vaccine.[5] In this paper, we use the probit model to show the existence of an inverse relationship between practice of MC and HIV prevalence in different regions of Kenya. This will help gain insights on important factors to be looked into in order to encourage male circumcision as an intervention for reduction of the spread of HIV/AIDS in the country.

## 1.2. Literature Review

[4] Using Decision Makers' Program Planning Tool (DMPPT) to model the impact of Voluntary Medical Male Circumcision (VMMC), a study based on 13 countries in eastern and southern Africa concluded that scaling up of VMMC will significantly reduce HIV infections in the countries and avoid HIV care costs, hence lower health system costs.

[6] A randomized controlled trial was carried out on 2784 men in Kisumu, Kenya. A group of men were assigned the intervention group (circumcision) and the other group was assigned a control group (delayed circumcision) and were assessed by HIV testing, medical examinations and behavioural interviews after a few months. The results showed that there's significant evidence that male circumcision reduces the risk of HIV acquisition in Kenya.

[7] The inner surface of the foreskin of a man's penis contains frenulum which is a common site for viral entry in primary HIV infection in men. Circumcision, which is the removal of this foreskin, especially at puberty, would be the most effective intervention in reduction of HIV transmission since it would be done before the men become sexually active.

[8] In a study on factors affecting HIV epidemic, backward multiple regression analysis was used to determine the major factors that are associated with HIV epidemic, from a set of five indicators. The indicators included; contraceptive prevalence rate, physicians' density, proportion of Muslim populations, adolescent fertility rate, and mean year of schooling. From the results, they were able to conclude that adolescent fertility rate significantly increases the rate of spread HIV infection, while the religious Muslim restrictions and availability of physicians decreases the rate of HIV infection in the society.

[9]The degree of prevention of acquisition of HIV infection that is provided by male circumcision can be equaled to that of prevention by a highly effective vaccine. Male circumcision may be an effective way of reducing HIV transmission in Sub-Saharan Africa.

[10] The rapid increase of HIV not only in Africa, Asia, but also in Europe, Mexico, and recently Central America, is a satisfying reason to an investment in MC as a part of a comprehensive HIV prevention package. The practice of MC could prevent millions of HIV new infections especially in sub-Saharan Africa and save on future treatment costs.

[11] In a study on HIV survey carried out in Kenya, male circumcision was identified as one of the factors associated with HIV prevalence. The study pointed that, based on

significant evidence; prevention efforts targeting behavioral and biologic factors like MC should be enhanced

[12] The Probit model is used in modelling dichotomous outcome variables. For this model, the link function is called the probit link. It uses the inverse of the cumulative distribution function of the standard normal distribution to transform probabilities to the standard normal variable.

[13] The use of the probit model was first used by Bliss in 1934. According to Greenberg, 1980, the aim of the study was to find an effective pesticide to control insects that fed on grape leaves. From the study, it was discovered that the response to a dose of pesticide was sigmoid. He applied the probit transformation to transform the sigmoid shape dose response curve to a linear relationship. According to some sources, probit analysis remains the preferred method in understanding dose-response relationship. The probit model assumes that random errors have a multivariate normal distribution. This model is deemed attractive because the normal distribution provides a good approximation to many other distributions.

## 2. Methodology

The study will make use of Kenya Aids Indicator Survey (KAIS), 2012 data obtained from Kenya National Bureau of Statistics (KNBS). The KAIS data was based on a study on adults and adolescents aged 15 to 64 and children aged 18 months to 14 years. About 14,000 women and men aged 15 to 64 years participated in the survey. The key objective of survey was to collect high quality data on the prevalence of HIV and Sexually Transmitted Infections (STI) among adults, and to assess knowledge of HIV and STI in the populations.

The study will involve two dependent variables, i.e. HIV and Male Circumcision (MC). The explanatory variables include age, residence, province, religion, and marital status. The bivariate probit model to be fitted is indicated below:

$$HIV = \beta_0 + \beta_1 Age + \beta_2 Residence + \beta_3 Province + \beta_4 Marital Status + \beta_5 Religion \quad (1)$$

$$MC = \beta_0 + \beta_1 Age + \beta_2 Residence + \beta_3 Province + \beta_4 Marital Status + \beta_5 Religion \quad (2)$$

### 2.1. Bivariate Probit Model

In the bivariate probit model, there are two binary dependent variables  $Y_1$  and  $Y_2$ .

The link function for the probit model is given by:

$$F(X'\beta) = \phi(X'\beta) = \int_{-\infty}^{x'\beta} \phi(z) dz \quad (3)$$

Let  $Y_1^*$  and  $Y_2^*$  be two latent variables. A latent variable is a hidden variable, meaning that it is not directly observed, but rather inferred through a mathematical model.

Each observed variable take on the value 1 if and only if its underlying continuous latent variable takes on a positive value:

$$Y_1 = \begin{cases} 1 & \text{if } Y_1^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$Y_2 = \begin{cases} 1 & \text{if } Y_2^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

With

$$Y_1^* = X_1\beta_1 + \varepsilon_1 \tag{6}$$

$$Y_2^* = X_2\beta_2 + \varepsilon_2 \tag{7}$$

and

$$\begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{Bmatrix} | X \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \rho & 1 \\ 1 & \rho \end{bmatrix} \right) \tag{8}$$

$$\begin{aligned} P(Y_i = 1) &= P(Y_i^* > 0) = \phi(X_i^T \beta) \\ &= \phi(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}) \end{aligned} \tag{9}$$

$$\begin{aligned} P(Y_i = 0) &= P(Y_i^* \leq 0) = 1 - \phi(X_i^T \beta) \\ &= 1 - \phi(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}) \end{aligned} \tag{10}$$

Fitting the bivariate model will involve estimating the values of  $\beta_1, \beta_2$  and  $\rho$  through maximization of the likelihood.

The Maximum Likelihood Estimation technique is used to estimate the probit model parameters. It focuses on choosing parameter estimates that gave the highest probability or likelihood of obtaining the observed values of the dependent variables. The likelihood is given as:

$$\begin{aligned} L(\beta_1, \beta_2) &= \prod \{ P(Y_1 = 1, Y_2 = 1 | \beta_1, \beta_2)^{Y_1 Y_2} * \\ &P(Y_1 = 0, Y_2 = 1 | \beta_1, \beta_2)^{(1-Y_1)Y_2} * \\ &P(Y_1 = 1, Y_2 = 0 | \beta_1, \beta_2)^{Y_1(1-Y_2)} * \\ &P(Y_1 = 0, Y_2 = 0 | \beta_1, \beta_2)^{(1-Y_1)(1-Y_2)} \} \end{aligned} \tag{11}$$

From the univariate model:

$$Y^* = X\beta + \varepsilon \tag{12}$$

$$Y_1 = \begin{cases} 1 & \text{if } Y_1^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

so for  $Y = 1$

$$\Rightarrow Y^* = X\beta + \varepsilon > 0 \tag{13}$$

$$\Rightarrow \varepsilon > -X\beta \tag{14}$$

$$\text{So } Y_1 = 1 \text{ is } \varepsilon_1 > -X_1\beta_1 \text{ and } Y_1 = 0 \text{ is } \varepsilon_1 < -X_1\beta_1 \tag{15}$$

and also if  $Y_2 = 1$  is  $\varepsilon_2 > -X_2\beta_2$

$$\text{and } Y_2 = 0 \text{ is } \varepsilon_2 < -X_2\beta_2 \tag{16}$$

Substituting the latent variables  $Y_1^*$  and  $Y_2^*$  in the probability functions and taking logs gives;

$$\begin{aligned} \sum & [Y_1 Y_2 \ln P(\varepsilon_1 > -X_1\beta_1, \varepsilon_2 > -X_2\beta_2) \\ & + (1 - Y_1) Y_2 \ln P(\varepsilon_1 < -X_1\beta_1, \varepsilon_2 > -X_2\beta_2) \\ & + Y_1 (1 - Y_2) \ln P(\varepsilon_1 > -X_1\beta_1, \varepsilon_2 < -X_2\beta_2) \\ & + (1 - Y_1) (1 - Y_2) \ln P(\varepsilon_1 < -X_1\beta_1, \varepsilon_2 < -X_2\beta_2)] \end{aligned} \tag{17}$$

Recall that:

$$Y^* = X\beta + \varepsilon \text{ when } Y = 1 \Rightarrow Y = \phi(X\beta) \tag{18}$$

$$\text{when } Y = 0 \Rightarrow Y = \phi(-X\beta) \tag{19}$$

$$\therefore Y_1 = \phi(X_1\beta_1) \tag{20}$$

$$Y_2 = \phi(X_2\beta_2) \tag{21}$$

So we will have:

$$\begin{aligned} & \sum [Y_1 Y_2 \ln \phi(X_1\beta_1, X_2\beta_2, \rho) + \\ & (1 - Y_1) Y_2 \ln \phi(-X_1\beta_1, X_2\beta_2, -\rho) + Y_1 (1 - \\ & Y_2) \ln \phi(X_1\beta_1, -X_2\beta_2, -\rho) + \\ & (1 - Y_1) (1 - Y_2) \ln \phi(-X_1\beta_1, -X_2\beta_2, \rho)] \end{aligned} \tag{22}$$

### 2.2. Marginal Probability Effect in the Probit Model

Marginal Effect determines the magnitude of change of the conditional probability of the outcome variable when you change the value of a regressor, holding all the regressors constant at some value.

$$\begin{aligned} & P(Y_i | X_{1i}, X_{2i}, \dots, X_{ki}; \beta_0, \beta_1, \dots, \beta_p) \\ & = \phi(\beta_0 + \sum_{k=1}^p \beta_k X_{ki}) \end{aligned} \tag{23}$$

$\phi(\cdot)$  is the cumulative distribution function of the standard normal distribution.

This means that conditional of the regressors, the probability that the outcome variable Y is 1, is a certain function of a linear combination of the regressor.

$$\frac{\partial P(Y_i=1 | X_{1i}, \dots, X_{ki}; \beta_0, \dots, \beta_p)}{\partial X_{ki}} = \beta_k \phi(\beta_0 + \sum_{k=1}^p \beta_k X_{ki}) \tag{24}$$

To evaluate or calculate the impact of  $X_k$  on, the researchers need to choose values for all other explanatory variables.

One approach would be to set all variables to their means or median.

Another approach would be to fix values of the other explanatory variables and let  $X_k$  vary from its minimum to maximum values, and then plot how the marginal effect of  $X_k$  changes across its observed range of values.

## 3. Results and Discussion

### 3.1. Introduction

In this section, a probit regression was applied separately for the dependent variables HIV and MC to determine the significant variables for the respective dependent variable. They were then regressed jointly as a bivariate probit model. The results of the estimated coefficient values (the z-scores), the probabilities and the p-value of the covariates are discussed in this section. The discussion was based on the findings. The level of significance used throughout this study was at 5%.

3.2. Univariate Probit Model for HIV

Table 1. Results for HIV prevalence in different age categories.

Age category	z-score	probability	p-value
15-19	0.0000	-	-
20-24	0.1301	0.5517	0.5650
25-29	0.7450	0.7719	0.0000
30-34	0.9494	0.8288	0.0000
35-39	0.9051	0.8173	0.0000
40-44	1.1290	0.8705	0.0000
45-49	1.0781	0.8595	0.0000
50-54	0.9756	0.8354	0.0000
55-59	0.6397	0.7388	0.0130
60-64	0.7772	0.7814	0.0030

From Table 1, the results showed that the age-groups 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59 and 60-64 are significant in explaining the incidence of HIV.

The positive z-scores indicate that occurrence of MC is more likely in the respective age-groups. At a 95% confidence interval; the study concluded that the age group ranging from 25-64 years were significant in determining the incidences of HIV in Kenya.

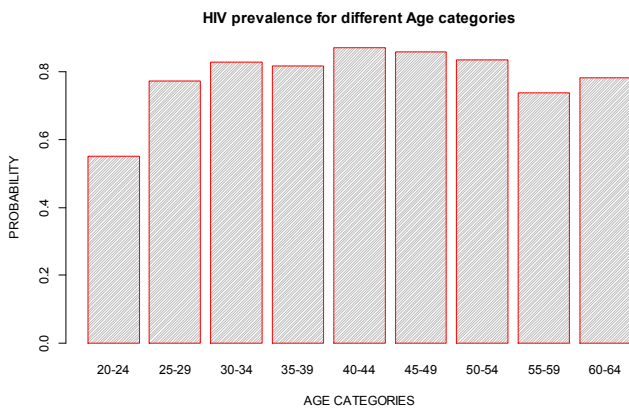


Figure 1. Probability of HIV prevalence in different age categories.

Table 2. Results for HIV prevalence by residence.

Residence	z-score	probability	p-value
Rural	0.0000	-	-
Urban	0.2877	0.6132	0.001

From Table 2, we see that urban residence is significant in explaining HIV. The positive z-score indicates that urban residence increases the likelihood of prevalence of HIV infection compared to rural areas.

Table 3. Results for HIV prevalence by Province.

Province	z-score	probability	p-value
North Eastern	0.0000	-	-
Central	-0.1977	0.4217	0.2480
Coast	-0.0270	0.4892	0.8650
Eastern	-0.2118	0.4161	0.1990
Nyanza	0.8907	0.8135	0.0000
Rift Valley	0.0389	0.5155	0.7890
Western	0.1241	0.5494	0.4540

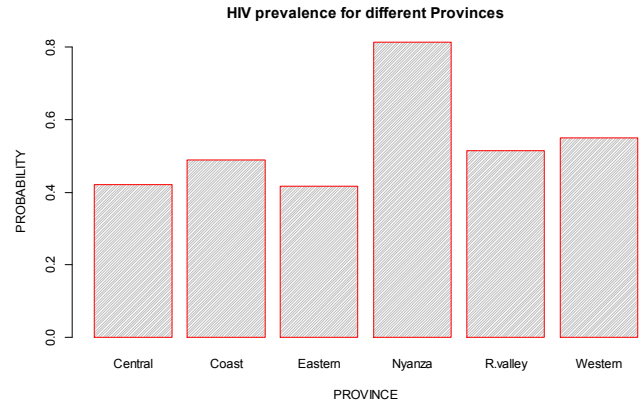


Figure 2. Probability of HIV prevalence in different provinces.

The results in Table 3 show that Nyanza province (p-value=0.000<0.05) is significant in explaining the incidence of HIV infection. The bar graph on figure 2 shows the high probability of HIV prevalence in the particular province. At a 95% confidence interval, the study noted that Nyanza significantly contributed to the national HIV statistics.

Table 4. Results for HIV prevalence by Marital Status.

Marital status	z-score	probability	p-value
Never Married	0.0000	-	-
Widowed	0.7280	0.7667	0.0100
Separated/Divorced	0.1572	0.5624	0.4290
Married/cohabiting	0.0516	0.5206	0.6770

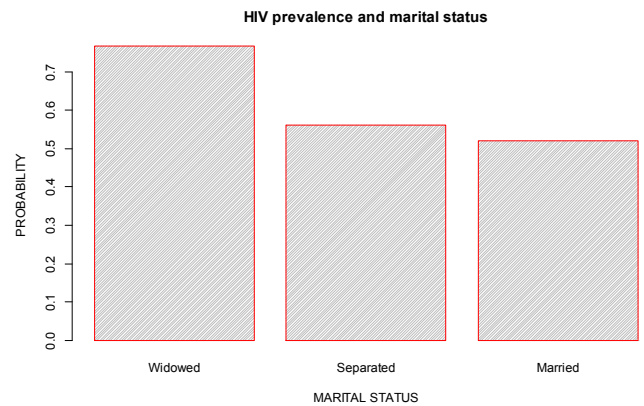


Figure 3. Probability of HIV prevalence in the different marital status categories.

Table 5. Results for HIV prevalence by Religion.

Religion	z-score	probability	p-value
Roman Catholic	0.0000	-	-
Protestant	-0.0534	0.4787	0.5340
Muslim	-0.2675	0.3945	0.1460
No Religion	-0.0104	0.4959	0.9620
Other	0.2223	0.5880	0.3730

Table 4 shows that being widowed was significant in determining the incidence of HIV infection. This is also reflected in Figure 3 that shows high prevalence of HIV among the widowed group. Thus at 95% confidence interval, the study concluded that widows and widowers had higher incidences of HIV in Kenya.

The results for religion as indicated in Table 5 showed that religion was not significant in determining the HIV incidences in Kenya. Muslim religion records the lowest probability of HIV prevalence, as shown in Figure 4.

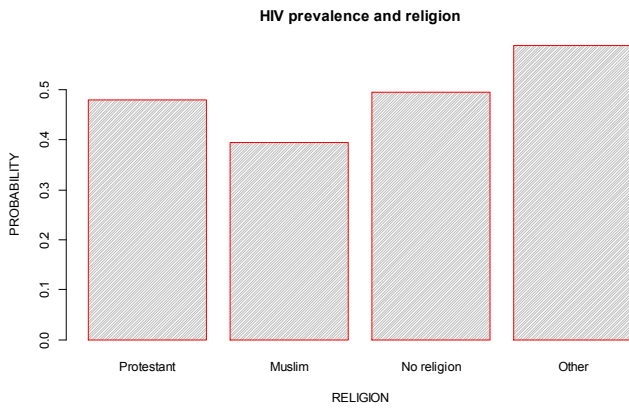


Figure 4. Probability of HIV prevalence in different religions.

### 3.3. Univariate Probit model for Male Circumcision (MC)

Table 6. Results for Univariate Probit model for Male Circumcision (MC).

Variable	MC		
	Z-Score	Probability	P-Value
<b>Age Category</b>			
15-19	0.0000	-	-
20-24	0.4991	0.6912	0.0000
25-29	0.6415	0.7394	0.0000
30-34	0.7307	0.7675	0.0000
35-39	0.7023	0.7588	0.0000
40-44	0.6261	0.7344	0.0000
45-49	0.4308	0.6667	0.0040
50-54	0.6723	0.7493	0.0000
55-59	0.6299	0.7356	0.0000
60-64	0.6533	0.7432	0.0000
<b>Residence</b>			
Rural	0.0000	-	-
Urban	-0.1601	0.4364	0.0140
<b>Province</b>			
North Eastern	0.0000	-	-
Central	0.4213	0.6632	0.0010
Coast	0.6422	0.7396	0.0000
Eastern	0.4049	0.6572	0.0000
Nyanza	-1.0129	0.1556	0.0000
Rift Valley	0.0178	0.5071	0.8590
Western	0.0304	0.5121	0.7870
<b>Marital Status</b>			
Never Married	0.0000	-	-
Widowed	-0.4971	0.3096	0.0710
Separated/Divorced	-0.3700	0.3557	0.0230
Married/cohabiting	-0.1530	0.4392	0.1120
<b>Religion</b>			
Roman Catholic	0.0000	-	-
Protestant	-0.0686	0.4727	0.2860
Muslim	0.6226	0.7332	0.0020
No Religion	-0.1183	0.4529	0.4180
Other	-0.9309	0.1759	0.0000

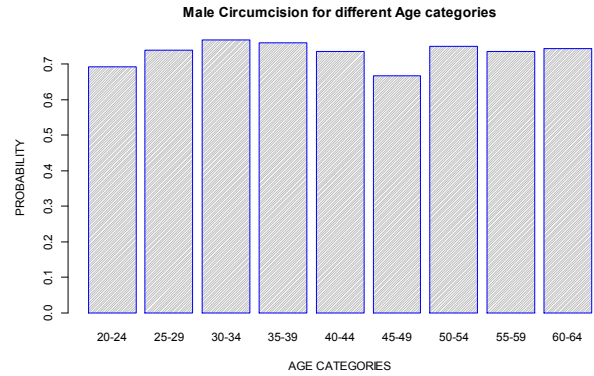


Figure 5. Practice of MC in different age categories.

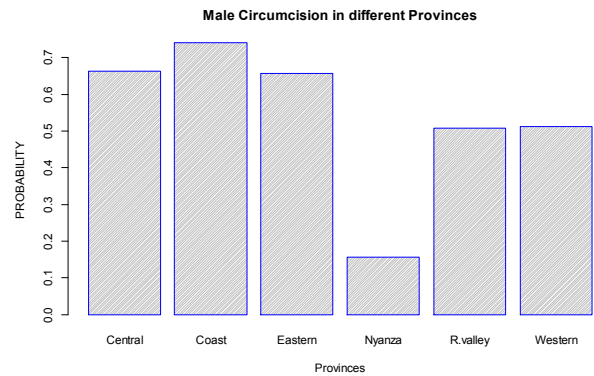


Figure 6. MC prevalence in different age categories.

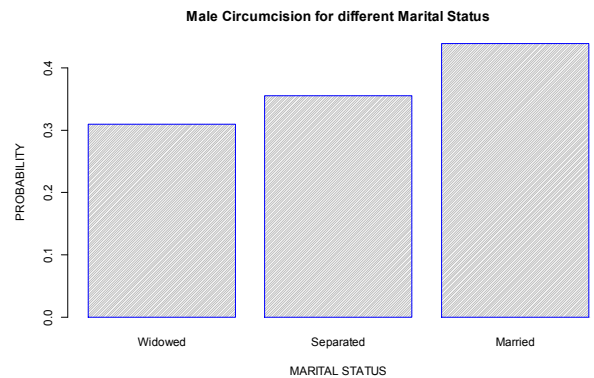


Figure 7. MC in different marital status.

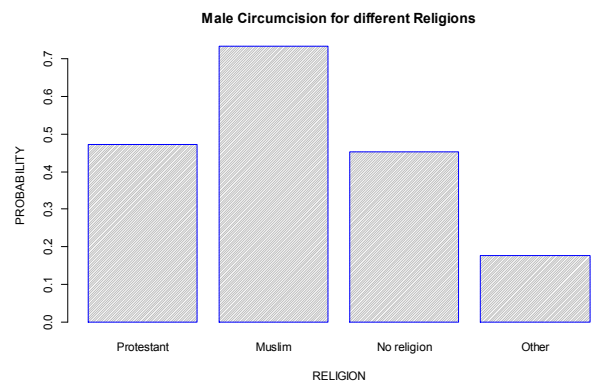


Figure 8. MC practice in different religions.

The study noted from the results in Table 6 that the age-group ranging from 20-24, 25-29, 30-34, 35-39, 40-44,

45-49, 50-54, 55-59 and 60-64 are significant in explaining male circumcision. The positive z-scores indicate that occurrence of MC is more likely in the respective age-groups. We also note that practice of MC in urban areas of Kenya is more likely compared to the rural areas. The negative z-score exhibited in Nyanza province indicates that the practice of MC is less likely to happen in the particular region.

The probit regression results also showed that Central, Eastern, Coast and Nyanza provinces are significant in explaining the occurrence of MC while Rift Valley and Western provinces are not. Widow/divorced persons and Muslim religion have a significant association with MC.

From figure 6, it is noted that practice of male circumcision is high in Coast, Eastern and Central provinces, while Nyanza province reports the lowest level of male circumcision.

Muslim religion also shows a high practice of male circumcision, as noted in Figure 8.

### 3.4. Bivariate Probit Model for Joint Dependent Variables

Table 7. Results for Bivariate Probit model.

Variable No.	Variable	HIV		MC	
		z-score	probability	z-score	probability
Age Category					
R.V	15-19	0.0000	-	0.0000	-
1	20-24	0.0348	0.5139	0.5375	0.7045
2	25-29	0.6419	0.7395	0.7227	0.7651
3	30-34	0.8396	0.7994	0.8860	0.8122
4	35-39	0.7977	0.7875	0.7593	0.7762
5	40-44	1.0091	0.8435	0.7122	0.7618
6	45-49	0.9672	0.8333	0.5725	0.7165
7	50-54	0.8632	0.8060	0.7592	0.7761
8	55-59	0.5444	0.7069	0.6696	0.7484
9	60-64	0.6728	0.7495	0.6909	0.7552
Province					
R.V	North Eastern	0.0000	-	0.0000	-
10	Central	-0.1817	0.4279	0.4043	0.6570
11	Coast	-0.0335	0.4866	0.6192	0.7321
12	Eastern	-0.1927	0.4236	0.3842	0.6496
13	Nyanza	0.8321	0.7973	-0.9643	0.1674
14	Rift Valley	0.0326	0.5130	0.0360	0.5144
15	Western	0.1016	0.5405	0.0128	0.5051
Residence					
R.V	Rural	0.0000	-	0.0000	-
16	Urban	0.2689	0.6060	-0.1727	0.4314
Marital Status					
R.V	Never Married	0.0000	-	0.0000	-
17	Widowed	0.7203	0.7643	-0.6970	0.2429
18	Separated/Divorced	0.1319	0.5525	-0.4288	0.3340
19	Married/cohabiting	0.0502	0.5200	-0.1803	0.4285
Religion					
R.V	Roman Catholic	0.0000	-	0.0000	-
20	Protestant	-0.0555	0.4779	-0.0880	0.4649
21	Muslim	-0.2754	0.3915	0.5774	0.7182
22	No Religion	-0.0222	0.4911	-0.1519	0.4396
23	Other	0.2822	0.6111	-0.7254	0.2341

R.V. = Reference Variable  
Observed rho value = -0.3872

### Cumulative Results

The results produced a negative correlation coefficient ( $\rho = -0.3872$ ), indicating an inverse relationship between HIV prevalence and male circumcision practice.

The graph below shows this effect;

Relationship between HIV and male circumcision

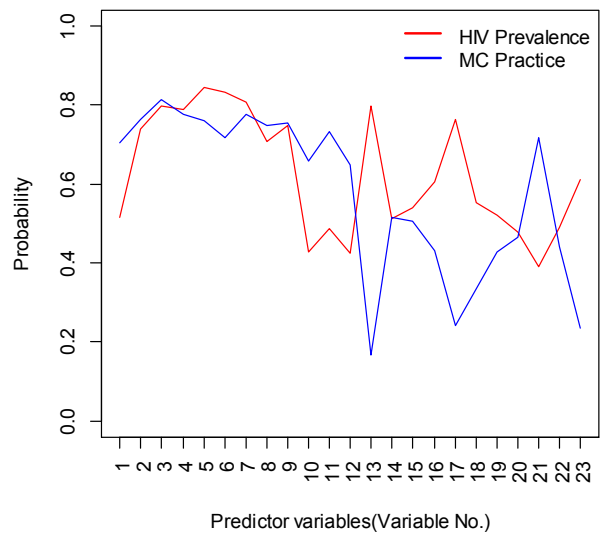


Figure 9. Graph of HIV and MC jointly regressed on all the predictor variables. The predictor variables are presented as variables numbers as per the table.

Age group 40-45 (variable no. 5) is noted to have the highest HIV prevalence in the Age Category. Nyanza province (variable no.13) showed a high prevalence of HIV infection, and low practice of MC. It can also be noted from the graph, that Muslim religion (variable no.21) has low prevalence of HIV. This can be attributed to the fact that many Muslim communities practice male circumcision before puberty, which can be the most immediate effective intervention for reducing HIV transmission since it is done before young men are likely to become sexually active. [7]

## 4. Conclusion and Recommendation

From the Univariate probit model for HIV, the p-values of the parameters estimates revealed a number of factors that are associated with HIV prevalence. Age 25-64 and urban residence have a significance in HIV prevalence. Nyanza province is well noted as the only province that is significantly associated with HIV prevalence. Apart from the widowed group, other categories of marital status and the religion of an individual are not associated with the prevalence of HIV.

From the Univariate probit model for MC, factors that are significant in explaining the practice of MC are: Age 20-64, urban residence, Provinces-Central, Coast, Eastern, and Nyanza, Separated /Divorced persons, and the Muslim religion.

The inverse relationship of HIV and MC is well brought out in Nyanza province, widowed group of people and the Muslim religion. In these categories, where there is high practice of

male circumcision, a low prevalence of HIV is recorded and vice versa. The result of a negative rho value ( $\rho = -0.3872$ ) when HIV and MC are modelled jointly, reveal the negative correlation, which leads to the conclusion that where male circumcision is highly practiced, low prevalence of HIV is noted. This in line with findings from several previous studies.

The findings from this study provide motivation for recommending male circumcision as a worth investment towards reduction of HIV prevalence in Kenya. The study can be done using different reference variables in order to see whether the results will be consistent.

---

## References

- [1] IFC International Finance Corporation. IFC against AIDS Partnerships lists: Kenya. 2004.
- [2] Nasr Ahmed "State of the Globe: The Relationship between Male Circumcision and Genitourinary Infections." *J Glob Infect Dis*, 2012.
- [3] N. Westercamp & R. C. Bailey "Acceptability of Male Circumcision for Prevention of HIV/AIDS in Sub-Saharan Africa: A Review" Springer Science+Business Media, LLC 2006.
- [4] Njeuhmeli E. et al. "Voluntary Medical Male Circumcision: Modeling the Impact and Cost of Expanding Male Circumcision for HIV Prevention in Eastern and Southern Africa." Ed. Stephanie L. Sansom. *PLoS Medicine* 8.11 (2011): e1001132. PMC. Web. 9 Nov. 2015.
- [5] Klausner, Jeffrey D et al. "Is Male Circumcision as Good as the HIV Vaccine We've Been Waiting For?" *Future HIV therapy* 2.1 (2008): 1–7. PMC. Web. 29 Oct. 2015.
- [6] Bailey, Robert C. Moses S, Parker CB, et al. "Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomized controlled trial." 643-656 *The Lancet* 369.9562 (2007).
- [7] Szabo, Robert, and Roger V Short. "How Does Male Circumcision Protect against HIV Infection?" *BMJ: British Medical Journal* 320.7249 (2000).
- [8] M. Mondal and M. Shitan "Factors affecting the HIV/AIDS epidemic: an ecological analysis of global data" *Afr Health Sci.* 13(2): 301–310. doi: 10.4314/ahs.v13i2.15; June 2013 June.
- [9] Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A "Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk" the ANRS 1265 Trial. *PLoS Med.* 2(11),E298 (2005).
- [10] Williams BG, Lloyd-Smith JO, Gouw E, Hankins C, Getz WM, Hargrove J, de Zoysa I, Dye C, Auvert B. The potential impact of male circumcision on HIV in Sub-Saharan Africa. *Plos Med* 2006.
- [11] Oluoch T. et al. "Correlates of HIV Infection Among Sexually Active Adults in Kenya: A National Population-Based Survey." *The Open AIDS Journal* 5 (2011): 125–134. PMC. Web. 9 Nov. 2015.
- [12] Introduction to SAS. UCLA: Statistical Consulting Group. From <http://www.ats.ucla.edu/stat/sas/notes2/> (accessed November 24, 2007).
- [13] Greenberg, B. G. (1980). "Chester I. Bliss, 1899-1979." *International Statistical Review / Revue Internationale de Statistique* 8(1): 135-136.