International Journal of Advanced Research in Biological Sciences
ISSN: 2348-8069
www.ijarbs.com
DOI: 10.22192/ijarbs Coden: IJARQG(USA) Volume 6, Issue 10 -2019

Research Article

DOI: http://dx.doi.org/10.22192/ijarbs.2019.06.10.009

Statistical Analysis of Factors that Affects Women's Knowledge on Mother-to-Child Transmission of HIV during Pregnancy in Ethiopia

Desalegn Dargaso Dana
Department of Statistics, College of Natural and Computational Science, Wolaita Sodo University, Wolaita Sodo, Ethiopia
***Corresponding author:** Desalegn Dargaso Dana
E-mail: *desdargaso@gmail.com*

Abstract

Human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) is one of the most encompassing and highly recognized infectious pandemics in our modern world today. Mother-to-child transmission of HIV is the major source of HIV infection among children under the age of fifteen years. The data were obtained from the 2016 Ethiopian Demographic and Health Survey with women in the age group 15-49 years. Binary logistic regression and descriptive statistical measures are applied. The results of the analysis showed that among the place of residence, urban women are more knowledgeable than rural women (72.3 and 67.1 percent, respectively) to report awareness about mother-to-child transmission of HIV during pregnancy. The Rural women are the least likely to know the facts about mother to child transmission of HIV compared with Urban women. The percent of knowing mother to child transmission of HIV increase as education level increase. According to the result of logistic regression, age, education level, religion, region, place of residence, frequency of reading newspaper, frequency of listening radio and HIV testing play significance role in determining women's knowledge on mother to child transmission of HIV. This study suggests that investment in women’s education should be a practical priority and the impact of education on knowledge about mother to child transmission of HIV increase with education level. Therefore, the government of Ethiopian should also consider and improve females' participant in school to reduce the prevalence of mother to child transmission of HIV. Based on the results of this study, the researcher found to be promising to address practical problems of mother to child transmission of HIV virus in Ethiopia. It can be learn from a study that in addition to the efforts being made to reduce the frequency of mother to child transmission of HIV in general and produce healthy children in the country.

Keywords: HIV/AIDS, pregnancy, statistical measures.

Introduction

Human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) is one of the most encompassing and highly recognized infectious pandemics in our modern world today. Not only has it changed the face of the world in terms of total overall deaths, but it also continues to contribute immensely to economic hardship, particularly in underdeveloped and developing countries. While HIV is transmit through sexual intercourse, the sharing of infected needles and blood products, and through vertical mother-to-child transmission during birth, currently no vaccinations or medications exist to eradicate the disease or to cure the disease after infection. Treatment is available in the form of antiretroviral medications that help to increase patients' life expectancy while they are living with HIV/AIDS. Some antiretroviral medications can prevent and eliminate the transmission of HIV from mother to child during pregnancy and birth. Unfortunately, access to those prenatal medicines is challenging, especially in resource-poor settings. Mother-to-child transmission is responsible for about 20% of all HIV infections, but is the route through which almost all pediatric infections occur. MTCT can occur during pregnancy, labor or breastfeeding. Advanced maternal disease, acute maternal infections during pregnancy and lactation, and co-morbidity with Sexually Transmitted diseases increase the risk of transmission [1].

Lack of awareness and misconceptions by the public about HIV/AIDS might have contributed to the spread of the disease. In addition, lack of access to information, education and counseling about the disease also contribute to continued high-risk behaviors. To date, the only protection against infection with the disease is to modify behavior pattern in order to minimize risk. Hence, information dissemination, education and communication continue to play a major role in the prevention of HIV/AIDS. Studies conducted in Ethiopia on knowledge, attitude and practice on various target groups provided a baseline data on sexual behavior, knowledge and attitude towards HIV/AIDS and the urgency of comprehensive health education as an intervention method. The transmission of HIV from a HIV-positive mother to her child during pregnancy, labor, delivery or breastfeeding is called mother-to-child transmission. In the absence of any intervention, transmission rates range from 15% - 45%. This rate can be reduced to below 5% with effective interventions during the periods of pregnancy, labor, delivery and breastfeeding. These interventions primarily involve antiretroviral treatment for the mother and a short course of antiretroviral drugs for the baby. They also include measures to prevent HIV acquisition in the pregnant woman and appropriate breastfeeding practices[2].

The reasons for an increasing mother to child transmission of HIV might include lack of knowledge of mothers on the risk of mother to child transmission, lack of access to provider health initiative and the benefits of preventive interventions, like antiretroviral drugs and infant feeding options. Mothers should be aware about HIV/AIDS and its routes of transmission to children and they should be motivated to recognize their status and be advised on primary and secondary preventive measures. However, the awareness of mothers on mother to child transmission of HIV and its prevention is not well studied in the counties. Pregnant women face many difficult decisions, including decisions around HIV testing, treatment options and infant feeding. A woman's male partner (s), extended family, greater community and health care setting all influence her decision and ability to take advantage of MTCT prevention. In the developing world, there is a lack of access to medications in general and antiretroviral drugs in particular. In addition, there is very little access to good health care for women both before and after birth, limited HIV counseling and testing and high stigma and discrimination against HIVPositive women [3]. Mother-to-child transmission of HIV is the spread of HIV from a woman living with HIV to her child during pregnancy, childbirth (also called labor and delivery), or breastfeeding (through breast milk). Mother-to-child transmission of HIV is also called perinatal transmission of HIV. The communication or transmission of HIV from HIV^+ mother to her child during pregnancy, delivery or breastfeeding is called mother-to-child communication/transmission of HIV. The main risk factor, which is also a barrier to the prevention of prenatal HIV transmission, is lack of awareness of HIV status among pregnant women. This study aims to identify the key factors that influence women's knowledge on mother-to-child transmission of HIV in Ethiopia to bear awareness regarding to vertical HIV transmission. The reasons for an increasing mother to child transmission of HIV might include lack of knowledge of mothers on the risk of mother to child transmission, lack of access to provider health initiative and the benefits of preventive interventions, like antiretroviral drugs and infant feeding options. Mothers should be aware about HIV/AIDS and its routes of transmission to children and they should be motivated to recognize their status and be advised on primary and secondary preventive measures. However, the awareness of mothers on transmission of HIV to child during pregnancy and its prevention is not well studied in Ethiopia. Therefore, this study is conduct to identify the determinant factors that affect women's knowledge on mother to child transmission of HIV/AIDS.

Methods

Source of data

The data used for this study was obtained from the 2016 Ethiopian Demographic and Health Survey ***[4]*** data conducted by Central Statistical Agency [5] in 2016 G.C. The principal objective of the 2011 Ethiopia Demographic and Health Survey (EDHS) is to provide current and reliable data on fertility and family planning behavior, child mortality, adult and maternal mortality, children's nutritional status, use of maternal and child health services, knowledge of HIV/AIDS, and prevalence of HIV/AIDS and anemia. An important feature of the data set is that it avails in-depth information on demographic and health aspects of households. The survey questionnaire included, among others respondent's knowledge about reproduction, contraception, pregnancy, postnatal care and breastfeeding, immunization and health, marriage, fertility preference, women's work status, maternal mortality, female circumcision, AIDS and other Sexually transmitted diseases , sexual activity.

Variables in the study

The dependent variable : is the dichotomous random variable
"HIV can't/can transmit mother-to-child during pregnancy period". "HIV can't transmit mother-to-child during pregnancy period" coded as 0 and "HIV can transmit mother-to-child during pregnancy period" coded as 1.

Independent variable: -The variables that are assumed to influence the prediction of "HIV can't/can transmit mother-to-child during pregnancy period" are presented below.

Table 3.1: Description of the variables and coding

S.N	Variable	Representation of Variable	Categories
1.	Age (age)	X_1	1 = 15-19 2 = 20-24 3 = 25-29 4 = 30-34 5 = 35-39 6 = 40-44 7 = 45-49
2.	Place of Residence (residence)	X_2	1 = Urban 2 = Rural
3.	Region (region)	X_3	1 = Tigray 2 = Affar 3 = Amahra 4 = Oromiya 5 = Somali 6 = Ben-Gumuz 7 = SNNP 8 = Gambela 9 = Harari 10 = Addis Ababa 11 = Dire Dawa
4.	Education level (education)	X_4	0 = No education 3 = Higher 2 = Secondary 1 = primary
5.	Frequency of listening radio (radio)	X_5	0 = not at all 1 = less than once a week 2 = at least once a week

6.	Marital status (marital)	X_6	0 = never married 1 = married/living together 2 = divorced/widowed
7.	Ever been tested for HIV (test)	X_7	0 = no 1 = yes
8.	Current pregnant (pregnant)	X_8	1 = yes 0 = no/not sure
9.	Religion (religion)	X_9	1 = Orthodox 2 = Catholic 3 = Protestant 4 = Muslim 5 = Traditional 6 = other
10.	Frequency of reading newspaper or magazine (news)	X_{10}	0 = Not at all 1 = Less than once a week 2 = At least once a week

Regression Analysis

Logistic regression can be binary or multinomial. The binary or Binomial logistic regression is the type of regression which is used when the dependent variable is a dichotomous and the independent variables are of any type while Multinomial logistic regression is used when the dependent variable has more than two categories. When multiple classes of the dependent variable can be ranked, then ordinal logistic regression is preferred to multinomial logistic regression. Continuous variables are not used as dependent variables in logistic regression.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 independent variables; to assess interaction effects; and to understand the impact of covariate control variables. The logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable (the natural log of the odds of the dependent variable occurring or not). In this way, logistic regression estimates the probability of a certain event occurring. Note that logistic regression calculates changes in the log odds of the dependent variable, not changes in the dependent variable itself as OLS regression does. However, logistic regression has many analogies to OLS regression: logit coefficients correspond to be coefficients in the logistic regression equation, the standardized logit coefficients correspond to beta weights, and a pseudo R-square statistic is available to summarize the strength of the relationship. Unlike OLS regression, however, logistic regression does not assume linearity of relationship between the independent variables and the dependent variable, does not require normally distributed variables, does not assume homoscedasticity, and in general has less stringent requirements [6]. The logistic regression is also preferred from multiple regression and discriminate analysis as it results in a biologically meaningful interpretation, it is mathematically flexible and easily used distribution and it requires fewer assumptions [7].

The Logistic Function

Consider the model

$$p = \alpha + \beta x \quad \ldots\ldots\ldots [1]$$

Where p is the probability of "success", α is the intercept of the line and β is its slope. This model is not feasible even on inspection. Since p is a probability, it is restricted to taking values between 0 and 1. The term $\alpha + \beta x$, in contrast, could easily yield a value that lies outside this range. We might try to solve this problem by fitting the model as

$$p = e^{\alpha+\beta x}$$

This equation/model guarantees merely the estimate of p is positive. Otherwise the term $e^{\alpha+\beta x}$although cannot be negative, it can result in a value that is greater than 1.To accommodate this final constraint, we fit a model of the form

$$p = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \quad \ldots\ldots\ldots [2]$$

The expression on the right, called a logistic function, cannot yield a value that is negative or greater than 1;consequently restricting the estimated value of p to the required range (between 0 and 1). If an event occurs with probability p, then the odds in favor of the event are $\frac{p}{1-p}$ to 1.Thus, if a success occurs with probability

$$p = \frac{e^{\alpha+\beta x}}{1 + e^{\alpha+\beta x}}$$

$$\frac{p}{1-p} = \frac{e^{\alpha+\beta x} / 1 + e^{\alpha+\beta x}}{1 / 1 + e^{\alpha+\beta x}}$$

$$= e^{\alpha+\beta x} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{[3]}$$

Taking the natural logarithm of each side of equation 3,
$\ln \frac{p}{1-p} = \ln_e \alpha + \beta x$, we obtain

$$= \alpha + \beta x \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{[4]}$$

Thus, modeling the probability p with logistic function is equivalent to fitting a linear regression model in which the continuous response y has been replaced by the logarithm of the odds of success for a dichotomous random variable. Instead of assuming that the relationship between p and x is linear, we assume that the relationship between $\ln \frac{p}{1-p}$ and x is linear [7]. The technique of fitting a model of this form is known as logistic regression.

The Multiple Logistic Regression Model

Consider a collection of k independent variables which will be denoted by the vector $X'=(x_1, x_2\ldots x_k)$. Let the conditional probability that the outcome is present be denoted by

$P(Y=1|X) = p(X)$. Then the logit of the multiple logistic regressions is given by the equation
$g(X) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$ and the odds in favor of success for the multivariate logistic regression will be

$$\ln \frac{p}{1-p} = \ln e^{g(x)}$$

$$= \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{[5]}$$

In which case $p(x) = \frac{e^{g(x)}}{1+e^{g(x)}}$

$$= \frac{e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k}}{1+e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k}} \quad (53)$$

Fitting the Logistic Regression Model

Suppose we have a sample of n independent observations of the pair (x_i, y_i), i=1, 2… n, where y_i denotes the value of the dichotomous outcome variable and x_i is the value of the independent variables for the i^{th} subject. Fitting the model requires that we obtain estimates of the values of α and β_1 (represented by a vector β').

In linear regression the method used most often for estimating unknown parameters is least squares. In that method we choose those values of β' which minimize the sum of squared deviations of the observed values of Y from the predicted values based upon the model. Under the usual assumptions for linear regressions the method of least squares yields estimators with a number of desirable statistical properties. Unfortunately when the method of least squares is applied to a model with a dichotomous outcome the estimators no longer have the same properties [7].

The general method of estimation that leads to the least squares function under the linear regression model (when the error terms are normally distributed) is called maximum likelihood. It is this method that provides the foundation for our approach to estimation with the logistic regression model. In a very general sense the method of maximum likelihood yields values for the unknown parameters which maximize the probability of obtaining the observed set of data. In order to apply this method we must first construct a function called likelihood function. The maximum likelihood estimators of these parameters are chosen to be those values which maximize this function. Thus, the resulting estimators are those which agree most closely with the observed data. If Y is coded as zero or one then the expression for p(x) above provides (for an arbitrary value of β', the vector of parameters) the conditional probability that Y is equal to 1 given x (i.e., $P(Y=1|x)$). It follows that 1- p(x) gives the conditional probability that Y is equal to 0 given x, $P(Y= 0|x)$. Thus, for those pairs (x_i, y_i), where $y_i=1$ the contribution to the likelihood function is $p(x_i)$, and for those pairs where $y_i=0$ the contribution to the likelihood function is 1-p (x_i), where the quantity $p(x_i)$ denotes the value of p(x) computed at x_i. A convenient way to express the contribution to the likelihood function for the pair (x_i, y_i) is through the Bernoulli distribution,

$\theta(xi) = p(xi)^{yi}[1 - p(xi)]^{1-yi}$ Where, i= 1, 2 …….n and $p(xi) = \frac{e^{\alpha+\beta xi}}{1+e^{\alpha+\beta xi}}$

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:

$$l(\beta) = \prod_{i=1} \theta(x_i) \quad \ldots\ldots (6)$$

The principle of maximum likelihood states that we use as our estimate of β the value which maximizes the expression in equation (6). However, it is easier mathematically to work with the log of equation (6). This expression, the log likelihood is defined as

$$L(\beta) = \ln[(\beta)] = \sum_{i=1}^{n} \{yi \ln[p(xi)] + (1 - yi)\ln[1 - p(xi)]\} \ldots\ldots [7]$$

To find the value of β that maximizes *equation (7)* we differentiate (β) with respect to α and β1 and set the resulting equation to zero. These equations are as follows:

$$\sum_{i=1}^{n} [yi - p(xi)] = 0 \ldots\ldots (a) and$$

$$\sum_{i=1}^{n} xi[yi - p(xi)] = 0 \ldots\ldots (b)$$

And these equations are called the likelihood equations. For logistic regression the expressions in (a) and (b) are non linear in α and β_1, and thus require special methods for their solution. These methods are iterative (like Newton Raphson) in nature and have been programmed into available logistic regression packages like SPSS, STATA, SAS etc. We can proceed to the multivariate case with the same steps followed in the univariate case above.

3.6 Model Building Strategies/Variable Selection

In modeling with many independent variables, one is usually concerned with the goal of selecting those variables that result in the "best" model within the scientific context of the problem. Having a basic plan to follow in selecting the variables for the model and assessing the adequacy of the model both in terms of the individual variables and from the point of view of the overall fit of the model is required for achieving this "best" model. It is also highlighted in [7] that successful modeling of a complex data set is part science, part statistical methods, and part experience and common sense.The traditional approach to statistical model building involves seeking the most parsimonious model that still explains the data. Recently researchers are shifting to including all scientifically relevant variables in the model, irrespective of their contribution to the model. This is based on the fact that it is possible for individual variables not to exhibit strong association while they do show considerable association when taken collectively. Both approaches have their merits and demerits and details of this can be found in [7].

In general the following steps are recommended by [7] to aid in the selection of variables for a logistic regression. Firstly, the selection process should begin with univariate analysis of each variable. Secondly, selection of variables for the multivariate analysis will follow based on the results in the univariate analysis along with all variables of known biologic importance. Finally, the importance of each variable included in the multivariate model should be verified by different model assessment techniques.

3.7.2 Hosmer- Lemeshow Test:-

The Hosmer –Lemeshow Test is another alternative in checking model fitness. This is based on the work of Hosmer who proposed grouping based on the values of the estimated probabilities [7]. The grouping can be either based on the percentiles of the estimated probabilities or fixed values of the estimated probabilities. In either case, the Hosmer – Lemeshow goodness of fit statistic, $C^{\wedge}$, is obtained by calculating the person chi-square statistic from the 2xg table of observed and expected frequencies and its formula is given as

$$\hat{c}=\sum_{k=1}^{g}\frac{(O_k-n_{k}p_k)^2}{n_{k}p_k(1-P_k)} \quad [10]$$

where g is the number of groups, n_k is the number of covariate patterns in the k^{th} group,

$Ok = \sum_{j=1}^{nk} yj$ is the number of responses among the n_k covariate patterns, and $pk = \sum_{j=1}^{nk} \frac{mjpj}{nk}$ is the average estimated probability. If the Hosmer-Lemeshow goodness-of-fit test statistic 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 are adequate to fit the data at an acceptable level. In other way the test statistic asymptotically follows a χ^2 distribution with *g-2* degrees of freedom. The number of risk groups may be adjusted depending on how many fitted risks are determined by the model.

Results and Discussion

The main objective of this study was to explain a range of potential factors which have a significant effect on HIV testing outcome of an individual in both urban and rural areas in Ethiopia by using 2016 Ethiopia Demographic and Health Survey data. Descriptive and binary logistic regression methods were used to measure the effect of determinants of women's knowledge on mother-to-child transmission of HIV during pregnancy period. The descriptive part of the data provides percentages of knowing the transmission of HIV during pregnancy time status of childbearing women and binary logistic regression analysis used to assess determinants that affect childbearing women's knowledge on mother-to-child transmission of HIV during pregnancy time and to predict the odds of knowing mother-to-child transmission of HIV in Ethiopia. The data are analyzed by using SPSS version 21 and STATA Version 15.

Summary of descriptive statistics

The result displayed in the table below (Table 4.1), illustrates that knowledge about mother-to-child transmission of HIV during pregnancy period is almost universal by background characteristics. About 67.5, 66.9,68.3, 67.3,69.4,71.0, and 64.3 percent of women in an age group 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 and 45-49 respectively have knowledge about mother to child transmission of HIV during pregnancy period and 32.5, 33.1, 31.7, 32.7, 30.6, 29.0 and 35.7 percent of women in an age group 15-19, 20-24, 25-29, 30-34, 35-39. 40-44 and 45-49 respectively have no knowledge about mother to child transmission of HIV during pregnancy period. Among regions, knowledge about mother to child transmission of HIV during pregnancy time ranges from 75% in Harari to 63.9% in Dire Dawa region. The level of awareness about mother to child transmission of HIV during pregnancy is notably lower in Dire Dawa region (63.9 percent of women) and higher in Addis Ababa (74.6 percent of women) than in other regions. Similarly, the sequence of having no awareness about mother to child transmission of HIV during pregnancy time among regional women from highest to lowest are explained as Harari (75%), Addis Ababa (74.6%), Amhara (69.3%), Somali (68.4%), Oromia (68.2%), Affar (67.7%), SNNPR (67.5%), Gambela (67.3%), BenshangulGumuz (66.2%), Tigray (64.5%) and Dire Dawa (63.9 %).

Concerning place of residence, urban women are more knowledgeable than rural women (72.3 and 67.1 percent, respectively) to report awareness about mother-to-child transmission of HIV during pregnancy. The Rural women are the least likely to know the facts about mother to child transmission of HIV compared with Urban women.

Table 4.1: Descriptive statistics

S.No.	Variables	Categories	Knowledge of mother-to-child transmission of HIV during pregnancy (%)	
			No	Yes
1.	Age in 5-year groups	15-19	32.5	67.5
		20-24	33.1	66.9
		25-29	31.7	68.3
		30-34	32.7	67.3
		35-39	30.6	69.4
		40-44	29.0	71.0
		45-49	35.7	64.3
2.	Region	Tigray	35.5	64.5
		Affar	32.3	67.7
		Amhara	30.7	69.3
		Oromia	31.8	68.2
		Somali	31.6	68.4
		Benishangul-Gumuz	33.8	66.2
		SNNPR	32.5	67.5
		Gambela	32.7	67.3
		Harari	25.0	75.0
		Addis Ababa	25.4	74.6
		Dire Dawa	36.1	63.9
3.	Type of place of residence	Urban	27.7	72.3
		Rural	32.9	67.1
4.	Highest educational level	No education	32.7	67.3
		Primary	30.3	69.7
		Secondary	32.3	67.7
		Higher	29.7	70.3
5.	Religion	Orthodox	31.8	68.2
		Catholic	31.9	68.1
		Protestant	31.6	68.4
		Muslim	32.1	67.9
		Traditional	33.3	66.7
		Other	34.1	65.9
6.	Frequency of reading newspaper or magazine	Not at all	32.0	68.0
		Less than once a week	30.6	69.4
		At least once a week	33.0	67.0
7.	Frequency of listening to radio	Not at all	32.2	67.8
		Less than once a week	30.5	69.5
		At least once a week	32.0	68.0
	Ever been tested for HIV	No	40.4	59.6
		Yes	25.6	74.4

8.	Currently pregnant	No or unsure	31.8	68.2
		Yes	33.2	66.8
9.	Current marital status	Never in union	30.6	69.4
		Married	32	68
		Living with partner	33.9	66.1
		Widowed	24.2	75.8
		Divorced	28.9	71.1
		No longer living together/separated	38.5	61.5

Logistic Regression Analysis

In univariate analysis, using Pearson chi square test (Table 4.2) displayed below, the variables that are found to be significant are Age, Educational background, marital Status, region, religion, frequency of reading newspaper, frequency of listening radio, pregnancy status, HIV test, place of residence, pregnant and marital status of women. The Pearson chi-square shows that all variables are selected as the candidates of multivariable logistic regression.
The test of hypothesis becomes;

H_0: There is no association between the outcome and predictor variable.
H_1: There is association between the outcome and predictor variable.

Table 4.2: The association between knowledge of mother to child transmission of HIV and each predictor

Variables	Pearson chi-square	D.f	P-value
Age in 5-year groups	43.05	6	0.030
Region	73.04	10	0.019
Type of place of residence	38.70	1	0.000
Highest Education level	55.05	3	0.000
Religion	62.25	5	0.000
Frequency of reading newspaper or magazine	49.90	2	0.034
Frequency of listening to radio	31.50	2	0.000
Currently Pregnant	4.60	1	0.000
Current marital status	9.62	5	0.000

Multivariate Finding

Multivariate analysis is done using the significant variables in the univariate analysis. Consequently, Age, Highest Educational level, current marital status, region, religion, frequency of reading newspaper, frequency of listening radio, pregnancy status and place of residence are included in the model as independent variables, but we have to identify potentially relevant variables for the model before analysis of logistic regression. To facilitate computation and interpretation, the coding scheme used in SPSS is given below in Table 4.3

Parameter estimation

Fitting a model to a set of data first entails estimating the unknown parameters in the model. We want to estimate the parameter βin such a way that the model fits the data best. There are different criteria to do this. The best known is maximum likelihood. The idea is to choose $\hat{\beta}$that maximizes the likelihood of the data.

Table 4.3: Estimates of the final logistic regression

Variables in the Equation								
	B	S.E.	Wald	Df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Age			3.717	6	.715			
15-19	.102	.261	.152	1	.697	1.107	.664	1.846
22-24	.081	.223	.130	1	.718	1.084	.700	1.679
25-29	.150	.220	.465	1	.495	1.161	.755	1.786
30-34	.108	.221	.238	1	.626	1.114	.722	1.718
35-39	.208	.224	.860	1	.354	1.231	.794	1.909
40-44	.276	.243	1.291	1	.256	1.318	.819	2.120
Region			15.818	10	.105			
Tigray	.056	.183	.093	1	.760	1.057	.739	1.513
Affar	.295	.160	3.420	1	.064	1.343	.982	1.836
Amhara	.294	.181	2.659	1	.103	1.342	.942	1.912
Oromia	.275	.161	2.932	1	.087	1.316	.961	1.803
Somali	.303	.150	4.096	1	.043	1.353	1.010	1.814
BenshengulGumuz	.170	.174	.959	1	.327	1.185	.844	1.665
SNNP	.188	.171	1.214	1	.270	1.207	.864	1.687
Gambela	.115	.191	.361	1	.548	1.122	.772	1.630
Harari	.577	.191	9.137	1	.003	1.781	1.225	2.589
Addis Ababa	.392	.207	3.582	1	.058	1.480	.986	2.220
Dire Dawa(ref)								
Residence(Urban)	.268	.102	6.884	1	.009	1.307	1.070	1.596
Rural(ref)								
Education	.079	.180	.191	1	.662	1.082	.761	1.539
No education	.177	.175	1.029	1	.310	1.194	.848	1.681
Primary	.029	.184	.025	1	.875	1.029	.718	1.476
Higher education								
Religion			.991	5	.963			
Orthodox	.067	.346	.038	1	.845	1.070	.543	2.107
Catholic	.059	.457	.017	1	.897	1.061	.433	2.601
Protestant	.114	.338	.115	1	.735	1.121	.578	2.174
Muslim	.008	.345	.001	1	.982	1.008	.513	1.980
Traditional	.129	.497	.068	1	.795	1.138	.430	3.015
Other(ref)								
News			11.727	2	.003			
Not at all	-.110	.098	1.259	1	.262	.895	.738	1.086
<Once a week	.112	.100	1.248	1	.264	1.119	.919	1.361
≥Once a week(ref)								
Radio Not at all			3.913	5	.562			
<Once a week	.323	.447	.521	1	.470	1.381	.575	3.317
≥Once a week(ref)	.336	.261	1.662	1	.197	1.399	.840	2.332
Radio Not at all	.190	.379	.251	1	.616	1.209	.575	2.542
<Once a week	.696	.397	3.075	1	.080	2.006	.921	4.366
≥Once a week(ref)	.460	.310	2.197	1	.138	1.583	.862	2.908
Constant	-.384	.562	.467	1	.494	.681		

The significance of the Wald statistic in the column with heading Sig indicates the importance of each predictor variable in the model and high values of the Wald statistic shows that the corresponding predictor variable is significant and important to model. The number in the odds ratio Exp(β), column of the logistic output are amounts by which the odds favoring y = 1 are multiplied, with each one unit increase in that x variables (if all other variables held fixed). Values of Exp (β) greater than one indicate that the corresponding variable increases the odds of being knowing Mother to Child transmission of HIV during pregnancy of individual women and values between 0 and 1 indicate a decrease in the odds of being knowing mother to child transmission of HIV during pregnancy of individual women. Prior to using the fitted model for the intended purpose, it is necessary to assess and diagnose the goodness of fit of the model.

Assessing the goodness of fit of the model

We begin our discussion of methods for assessing the fit of an estimated logistic regression model with the assumption that we are at least preliminarily satisfied with our efforts at the model building stage. By this, we mean that, to the best of our knowledge, the model contains those variables (main effects) that should be in the model and that variables have been entered in the correct functional form. Now we would like to know how effectively the model we have describes the outcome variable. This is referred to as its goodness-of-fit.

When the model building stage has been completed, a series of logical steps may be used to assess the fit of the model. The components of the proposed approach are (1) computation and evaluation of overall measures of fit, (2) examination of the individual components of the summary statistics, and (3) examination of other measures of the difference or distance between the components of y and $\hat{y}$.

The Hosmer-Lemeshow tests

Another commonly used test of model fit is the Hosmer and Lemeshow's goodness-of-fit test. The idea behind the Hosmer and Lemeshow's goodness-of-fit test is that the predicted frequency and observed frequency should match closely, and that the more closely they match, the better the fit. The Hosmer-Lemeshow goodness-of-fit statistic is computed as the Pearson chi-square from the contingency table of observed frequencies and expected frequencies. The hypothesis to be tested is,

H_0: Model fits the data versus H_1: Model does not fit the data.

Table 4.3:Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	1.915	8	.984

From the above, the Hosmer-Lemeshow test result reported chi-square value of 1.915 with p-value of 0.984 on 8 degrees of freedom. But this p-value is greater than the 0.10 and 0.05 levels showing that there is no difference between the observed and the model predicted values and hence estimates of the model fit the data at an acceptable level. Assessment of the interaction terms showed that none of them were statistically significant and hence were excluded from the model.

Recommendation

Based on the results of this study, the researcher found to be promising to address practical problems of mother to child transmission of HIV virus in Ethiopia. It can be learn from a study that in addition to theefforts being made to reduce the frequency of mother to child transmission of HIV in general and produce healthy children in the country; specific attention should be given on knowledge about mother to child transmission of HIV by taking the following in to consideration.

- Health workers need to be given continuous medical education regarding mother to child transmission of HIV since they are the main sources of information for rural and urban women.
- It recommends that investment in women's education should be a practical priority and the impact of education on knowledge about mother to child transmission of HIV increase with education level.
- More health workers should have trained to give awareness of mother to child transmission of HIV services.
- Ethiopian government should also invest heavily on women and additional study is required on the area of knowledge about mother to child transmission of HIV by considering detailand accurate information on various variables.

References

1. **UNAIDS (2009).** A strategic approach: HIV and AIDS and education (2009).
2. **WHO**: Mother to child transmission of HIV. Geneva, Switzerland: World Health Organization; 2015.
3. **UNAIDS (2010).** A strategic approach: HIV and AIDS and education (2009).
4. CSA (2016). Central Statistical Agency (2016).
5. EDHS (2016). Ethiopian Demographic and Health Survey (2016).
6. Stat Notes. Logistic Regression 2006.WWW 2. Chau. Nclu. Edu/ garson /PA765 /Logistic.
7. Hosmer D. and Lemeshow S.(1989). Applied Logistic regression .John Wiley and Snns. Inc .NEW York.

Access this Article in Online	
	Website: www.ijarbs.com
Quick Response Code	Subject: Epidemiology
DOI:10.22192/ijarbs.2019.06.10.009	

How to cite this article:
Desalegn Dargaso Dana. (2019). Statistical Analysis of Factors that Affects Women's Knowledge on Mother-to-Child Transmission of HIV during Pregnancy in Ethiopia. Int. J. Adv. Res. Biol. Sci. 6(10): 89-100.
DOI: http://dx.doi.org/10.22192/ijarbs.2019.06.10.009