ASSOCIATED RISK FACTORS OF HIV STATUS IN ETHIOPIA: MULTILEVEL ANALYSIS

Setegn Bayabil[1*], Solomon Sisay Mulugeta[1], Selamawit Getachew Wassihun[2]
Alebachew Taye Belay[3]
[1]Department of Statistics, Debre Tabor University, Debre Tabor, Amhara, Ethiopia
Corresponding Author: setegnbayabil@gmail.com
[2]Department of Statistics, Mekidela Amba University, Amhara, Ethiopia
[3]Department of Biostatistics, Debre Tabor University, Debre Tabor, Amhara, Ethiopia

Abstract

Background: HIV is a virus spread through certain body fluids that attacks the body's immune system, specifically the CD4 cells. In Ethiopia, the general trend of the prevalence rate varies across each year and in each region. The purpose of the current study was to identify the most risk factors Related to the HIV status of individuals in Ethiopia.

Methodology: The data for this study were taken from Ethiopia Demographic and Health Survey conducted in 2016 (EDHS 2016). Voluntarily HIV testing interviewers collected finger-prick blood specimens from women age 15-49 and men age 15-59 who consented to HIV testing. The analysis was conducted by binary logistic regression and multilevel logistic regression model.

Results: Among 26692 individuals included in the study, 55.1 % (14719) and 44.9 % (11973) were female and male individuals respectively. The prevalence of HIV/AIDS infection among VCT individuals was estimated to be 1.5% and varied by demographic, socio-economic, HIV Awareness, and sexual behavior factors of individuals. This study revealed that HIV awareness, sex, age at first sex, region, age, educational level, and place of residence were significant determinants of the HIV status of an individual. The odds of being HIV infected for a female individual was about 1.938 times the odds of male individuals and it implies that statistically significant. The odd of Individuals found in the Gambella region was 13.35 times being HIV positive than the odd of Individuals in the Tigray region.

Conclusion: The regional variation in the HIV status of an individual is statistically significant. In addition, HIV Awareness, sex, residence, educational level, and age at first sex were also found to be significant determinants of variation in HIV status of individuals among regions.

Keywords: HIV/AIDS, multilevel, binary logistic, Ethiopia, EDHS, VCT.

Background

HIV (human immunodeficiency virus) is a virus that assaults the body's immune system, primarily the CD4 cells, also known as T helper cells, and is disseminated through certain bodily fluids. HIV can damage so many of these cells over time that the body is unable to fight infections and disease. HIV and AIDS are not the same thing. HIV is an infection that damages the immune system and leads to AIDS. AIDS is a syndrome that develops when a person's immune system is too weak to fend off several infections, and it occurs after the HIV infection has progressed significantly. It is the last stage of HIV infection, when the body no longer has the ability to protect itself. There is currently no cure for HIV or AIDS. However, with the proper care and support[1]. HIV is a sexually transmitted infection, but it can also be spread by bodily fluid contact, such as sexual fluid exchange, blood, mother to child during pregnancy or childbirth, or breastfeeding. Non-sexual transmission can also occur when injection equipment, such as needles, is shared. They believed that the heterosexual epidemic could be limited if prevention efforts were focused on persuading the so-called core transmitters, such as sex workers and truck drivers, who are known to have several sex partners, to use condoms. For more than a decade, these obstacles slowed prevention campaigns in many nations [2]. Worldwide, HIV/AIDS poses an enormous challenge to the survival of mankind. The HIV pandemic remains the most serious of infectious disease challenges to public health. Promising developments have been seen in recent years in global efforts to address the AIDS epidemic, including increased access to effective treatment and prevention programs. According to [3], an estimated 36.9 million people globally are living with HIV of which 25.8 million are in Sub-Saharan Africa. There were two million new infections in 2014 with 1.4 million (65%) of these occurring in Sub-Saharan Africa. In 2019, around 1.2 million people died of AIDs related illnesses globally of which 795,890 were in Sub-Saharan Africa. As of June 2020, 15.9 million people living with HIV were

doi.org/10.31838/hiv22.01.05 Received: 02.07.21, Revised: 05.12.21, Accepted: 11.01.22

accessing antiretroviral therapy, up from 13.6 million in June 2019[3].

In Ethiopia, the prevalence rate fluctuates from year to year and from region to region. A trend analysis of the country's HIV/AIDS prevalence rate from 1982 to 2011 shows a steady growth until the late 1990s, then a steady drop in the years after 2000. The adult HIV prevalence rate in the United States was projected to be 0.2 percent in 1985, 3.2 percent in 1995, and 1.4 percent in 2005. However, according to the 2011 EDHS report, the prevalence rate in 2011 increased by a very small amount when compared to the prevalence rate in 2005. The outbreak is highly variable by region, with the lowest rates (0.9 percent) in SNNPR and the highest rates (6.5 percent) in Gambel [4].

HIV testing is an essential gateway to HIV prevention, treatment, care, and support services. In a study done in the case of Organization for Social Services for AIDS (OSSA) free standing of VCT center in urban Hawassa had a total of 1022 clients whose age were greater than or equal 15 years were included in the study, of which 57.7% were male and the rest 42.3% were females. Most of the clients were never married and concerning educational level, 56.0% were grade of 7-12 followed by grade 1-6, 19.7% and the rest 10.9% were belongs to tertiary and illiterate group respectively. About occupation, the majority of clients 20.5% were from students, 12.7% were from the merchant, 10.2% were from elementary occupation/unskilled, 4% from professional and the rest 52.5% were from other (housewife, vehicle driver, etc.) category. HIV/AIDS is a major public health concern and cause of death in many parts of Africa. Although the continent is home to about 15.2 % of the world's population, more than two-thirds of the total infected worldwide some 35 million people were Africans, of whom 15 million have already died [2].

The region hardest afflicted by HIV is East and Southern Africa. It is home to 6.2 percent of the world's population, but more than half of the world's HIV-positive persons (19.4 million people). There were 790,000 new HIV infections in 2016. Between 2010 and 2016, the number of new HIV infections among children aged 15 to 19 years dropped by 56%. (0-14 years). Adult new infections decreased by 29% over the same time period, albeit there was significant variance between nations. Mozambique, Uganda, and Zimbabwe experienced the steepest declines. The annual number of new infections increased in Ethiopia and Madagascar. In the region, women account for 56 percent of HIV-positive adults. Women in their twenties and thirties (15–24 years old) [5].

Ethiopia has made progress in reducing the number of people living with HIV/AIDS in the country, however the observed reductions remain insufficient in comparison to the targeted goals of the epidemic response. Given the size of the population and the magnitude of the damage, it will take several years for significant declines in HIV prevalence and incidence to be seen with concerted and sustained efforts. There are also advances in the availability, accessibility, and utilization of HIV/AIDS prevention, care, support, and treatment services, as well as improvements in the management of the epidemic and increasing resource availability. Therefore, this study is attempted to investigate the effect of individual socioeconomic, HIV/AIDS awareness, and sexual behavior factors in explaining associated risk factors of HIV status across regions in Ethiopia.

Methods

Study Setting

The source of data for this study was the 2016 Ethiopia Demographic and Health Survey (EDHS) which was collected by Central Statistical Agency (CSA). HIV testing Interviewers collected finger-prick blood specimens from women age 15-49 and men age 15-59 who consented to HIV testing. The outcomes of the survey at the national level, for the nine regional states and two city administrations of Ethiopia. Information can be used for various purposes, including program planning and evaluation. In this study, we have considered individuals who had HIV status.

Study design

The Demographic and Health Survey (DHS) is a periodic cross-sectional survey administered at the household level. By taking a blood sample for HIV testing from representative samples of the population of men and women in a country, the DHS can provide nationally representative estimates of HIV rates. However, it is important to recognize that population-based testing is dependent on the Population's willingness to be voluntarily tested for HIV.

Variables Under Investigation

Response Variables

The response variable is the dichotomous variable "HIV status of the respondent" (HIV positive=1, HIV negative=0).

Explanatory Variables

The covariates associated with HIV status of individuals in Ethiopia were categorized as Background individual /household socioeconomic and demographic characteristic includes age, urban/ rural residence, education, sex, region, and wealth index; HIV/AIDS awareness that is general awareness, modes of HIV transmission, and ways of preventing infection. HIV/AIDS awareness can be classified as No and Yes, And Sexual behavior factors include marital union status, age at first sex, and multiple sexual partners [4].

Statistical modeling

The current study was analyzed using binary logistic regression; to identify associated risk factors on the HIV status of an individual in Ethiopia and finally,

we assessed the effect of determinant factors and regional differences on the HIV status of individuals using a multilevel logistic regression model.

Logistic regression can be binomial, ordinal or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for the dependent variable in logistic regression is usually dichotomous, that is, "0" and "1" (which may represent, HIV positive vs negative or "success" vs "failure"). The dependent variable can take the value 1 with a probability of success π or the value 0 with a probability of failure 1- π, this type of variable is called a Bernoulli (or binary) variable. The logistic regression is a mathematically flexible and easily used distribution and it requires fewer assumptions[6].

Multilevel modeling is an approach that can be used to handle clustered or grouped data. It is also known as the hierarchical linear model. These models are particularly appropriate for research designs where data for participants are organized at more than one level (i.e., nested data). The units of analysis are usually individuals (at a lower level) who are nested within contextual/aggregate units (at a higher level)[7] In multilevel research, the structure of the data in the population is hierarchical, and a sample from such a population can be viewed as a multistage sample. First, we take a sample of units from the higher level (in our case from the 9 regions and two city administrations), and next, we take a sample subunit from available units (HIV status of individuals). Hierarchical models are statistical models that can be used to analyze nested sources of variability in hierarchical data, taking account of the variability associated with each level of the hierarchy [8].

The multilevel binary logistic regression model has a binary outcome (For individual HIV status). Let y_{ij} be the binary outcome variable, coded '0' or '1', associated with level one unit i nested within level two-unit j. In this study two-level model is used, level-1 is individual and level-2 is the regional level. The clustering of the data points within geographical regions offers a natural 2-level hierarchical structure of the data, i.e. individuals who had HIV status are nested within regions. We further simplify the presentation by assuming there is the individual-level predictor and regional-level factor. To provide a familiar starting point, we first consider a two-level model for binary outcomes with a single explanatory variable. Suppose we have data consisting of individuals who had HIV status (level one) grouped into regions (level two). Let Y_{ij} be the binary response for HIV status of individual i in region j and X_{ij}, an explanatory variable at the individual level. We define the probability of the response equal to one as π_{ij} = Pr (y_{ij} = 1) and let π_{ij} be modeled using a logit link function.

The intra-class correlation coefficient (ICC) measures the proportion of variance in the outcome explained by the grouping structure. It is an indication of the proportion of variance at the second level (region) and it can also be interpreted as the expected (population) correlation between two randomly chosen individuals within the same group.

The ICC is then calculated based on the following formula:

$$\text{ICC} = \frac{\sigma^2_{uo}}{\sigma^2_{uo}+\sigma_{e^2}}$$

Where σ_{e^2} is the variance of individual (lower) level units. Since the logistic distribution for the level, one residual variance implies a variance of $\sigma_e^2 = \frac{\pi^2}{3} = 3.29$ [8] and this formula can be rewritten as:

$$\text{ICC} = \frac{\sigma^2_{uo}}{\sigma^2_{uo}+3.29}$$

The ICC may range from 0 to 1. ICC = 0 indicates perfect independence of residuals: The observations do not depend on cluster membership. The chance of HIV-positive individuals does not differ from one region to another (there is no between-regional variation). When the ICC is not different from zero or negligible, one could consider running traditional one-level regression analysis[9]. However, ICC = 1 indicates perfect interdependence of residuals: The observations only vary between clusters.

Estimation of the Parameter

The most common methods for estimating multilevel logistic models are based on likelihood. Among the methods, Marginal Quasi Likelihood or MQL[[10],[11], and Penalized Quasi Likelihood or PQL[12] are the two prevailing approximation procedures. Both MQL and PQL are based on Taylor series expansion to achieve the approximation. Based on the usage of the first and second term of Taylor expansion, MQL and PQL are often known as first-order MQL and second-order MQL, first-order PQL, and second-order MQL respectively. After applying these quasi-likelihood methods, the model is then estimated using iterative generalized least squares (IGLS) or reweighted IGLS (RIGLS) [10].

Results

A total of 26,692 individuals who received VCT services from all regions were included in this study. The prevalence of HIV infection varies by the Socioeconomic and Demographic Characteristics of the VCT individuals. Table 1 shows that the distribution of Sero-Positive individuals by some demographic/HIV-related risk behavior variables. Sero-Positivity (HIV infection) is high among females than males. Among those clients, 55.1% were females and the rest 44.9% were males. Similarly, in Ethiopia, 1.1% of the male and 1.9% of the female clients were found to be HIV positive.

The majority of individuals, 69.5%, were from rural but the prevalence of HIV in Ethiopia is higher in urban (3.3%) areas than in rural areas (0.7%). In relation to age at first sex, 43.4% of the respondent less than 18 years and the result shows individuals greater than the 35-year age group was highly affected by the Virus (2.0%), and age group 25-34(1.7%) revealed the highest proportion of HIV. Among the respondents in this study, Most of them are Richest (47.9%), followed by Poorer and Middle (38.4% and 13.6% respectively). About Educational level, 36.6% were No education, and the remaining 38.7% and 24.7% belong to Primary, secondary, and above educational level respectively. HIV Prevalence of individuals who attended less educational level (No education or preschool) was lower than individuals who attended primary, secondary, and Higher education levels. This study was covered from Amhara, Oromia, SNNPR, Tigray, Addis Ababa, Somali, Benishangul-Gumuz, (13.2%, 12.9%, 12.5%, 11.1%, 10.1%, 8.1%, and 7.4%) respectively. Among regions, HIV prevalence is highest in Gambella (4.3%), Addis Ababa (3.4%), and Dire Dawa (2.2%). Among the respondents in this study were asked about their knowledge of ways of avoiding HIV infection as well as awareness of modes of transmission, 3.5% of the respondent does not have an awareness about HIV and 96.5% have an awareness.

Table 1: Univariate Chi-square analysis with response variable Risk Factors of HIV status among individuals in Ethiopia

Variable	Category	N (%)	HIV status of respondents		d.f	Chisquare	P-value
			HIV Negative	HIV Positive			
Region	Tigray	2966(11.1)	99.0%	1.0%	10	237.960	≤ 0.001*
	Afar	1783(6.7)	98.8%	1.2%			
	Amhara	3518(13.2)	98.8%	1.2%			
	Oromia	3454(12.9)	99.2%	0.8%			
	Somalia	2169(8.1)	99.7%	0.3%			
	Benishangual Gumuz	1969(7.4)	98.9%	1.1%			
	SNNPR	3340(12.5)	99.5%	0.5%			
	Gambela	1853(6.9)	95.6%	4.4%			
	Harari	1265(4.7)	97.4%	2.6%			
	Addis Abeba	2695(10.1)	96.6%	3.4%			
	Dire Dawa	1680(6.3)	97.8%	2.2%			
Place of Residence	Urban	8151 (30.5)	96.7%	3.3%	1	244.968	≤ 0.001*
	Rular	18541(69.5)	99.3%	0.7%			
Marital Status	Never Married	8522(30.8)	98.4%	1.6%	2	66.58	0.542
	Married	16172(64.4)	98.5%	1.5%			
	Separated	2059(4.8)	98.1%	1.9%			
Education level	No education Preschool	9775(36.6)	99.0%	1.0%	2	35.588	≤ 0.001*
	Primary	10327(38.7)	98.3%	1.7%			
	Secondary and above	6590(24.7)	98.0%	2.0%			
Sex	Male	11973(44.9)	98.9%	1.1%	1	31.571	≤ 0.001*
	Female	14719(55.1)	98.1%	1.9%			
Age	< 15 years	54(0.2)	98.1%	1.9%	3	61.361	≤ 0.001*
	15-24	7490(28.1)	99.4%	0.6%			
	25-34	9438(35.4)	98.3%	1.7%			
	Greater or equal to 35	9710(36.4)	98.0%	2.0%			
HIV Awareness	No	930(3.5)	99.5%	0.5%	1	6.272	0.012*
	Yes	25762(96.5)	98.4%	1.6%			
Age at first sex	Not had sex	5253(19.7)	99.0%	1.0%	4	19.673	≤ 0.001*
	<18 years	4790(17.9)	98.2%	1.8%			
	18-24	13568(50.8)	98.3%	1.7%			
	25-34	3010(11.3)	98.9%	1.1%			
	Greater or equal to 35 year	71(0.3)	97.2%	2.8%			
Wealth Index	Poorer	10263(38.4)	99.2%	0.8%	2	113.087	≤ 0.001*
	Middle	3638(13.6)	99.2%	0.8%			
	Richer	12791(47.9)	97.6%	2.4%			
Multiple sex Partner	more than one sex partner within the 12 months preceding the survey	1046(3.9)	98.5%	1.5%	1	0.000	0.998
	Otherwise	25646(96.1)	98.5%	1.5%			

Binary logistic regressions are fitted using the predictor variables, which are found to be significant in the univariate analysis. Based on results displayed in Table 1 and those predictor variables that are associated factors of HIV at 20% level significance is selected for multiple logistic regression analysis [13].

The result displayed in Table 2 showed that the predictor variables Region, place of residence, educational level, sex, age, HIV awareness, wealth, and Age at first sex were significantly associated with HIV infection at univariate analysis.

Table 2: Results of Binary Logistic Regression: - Estimates, Standard errors, Wald, Degree of freedom, P-values, estimated Odds-Ratio and Confidence Intervals

Variable	Category	$(\hat{\beta})$	S.E$(\hat{\beta})$	Wald	d.f	Sig	exp $(\hat{\beta})$	95% CI	
								lower	upper
HIV Awareness	0=Yes(ref.cat)	-	-	-	-	-	-	-	-
	1=No	0.913	0.461	3.914	1	0.048	2.491	1.009	6.153
Region				98.417	10	≤ 0.001*			
	1=Somali(ref.cat)	-	-	-	-	-	-	-	-
	2=Afar	1.543	0.465	11.004	1	0.001*	4.679	1.880	11.644
	3=Amhara	1.548	0.442	2.036	1	≤ 0.001*	4.703	1.978	11.182
	4=Oromia	1.209	0.454	0.002	1	0.008	3.351	1.376	8.165
	5=Tigray	1.199	0.452	7.022	1	0.008	3.316	1.366	8.049
	6=Benishangul gumuz	1.430	0.470	0.627	1	0.002	4.179	1.662	10.509
	7=SNNPR	0.752	0.477	2.175	1	0.115	2.121	0.833	5.398
	8=Gambela	2.591	0.433	37.341	1	≤ 0.001*	13.349	5.712	31.196
	9=Harare	1.748	0.451	4.437	1	≤ 0.001*	5.713	2.374	13.895
	10=Addis Abeba	1.552	0.431	2.512	1	≤ 0.001*	4.722	2.029	10.989
	11=Dire Dawa	1.344	0.447	0.324	1	0.003	3.836	1.597	9.213
Age				87.798	3	≤ 0.001*			
	0=< 15 year	-0.187	1.033	0.033	1	0.857	0.830	0.110	6.281
	1=15-24	-1.649	0.177	86.842	1	≤ 0.001*	0.192	0.136	0.272
	2=25-34	-0.427	0.111	14.842	1	≤ 0.001*	0.652	0.525	0.811
	3=greater than or equal to 35 year(ref.cat)	-	-	-	-	-	-	-	-
Sex	0=male(ref.cat)	-	-	-	-	-	-	-	-
	1=female	0.662	0.112	34.731	1	≤ 0.001*	1.938	1.555	2.415
Education level				19.884	2	≤ 0.001*			
	0=No education,p preschool(ref.cat)	-	-	-	-	-	-	-	-
	1=primary	0.528	0.140	14.146	1	≤ 0.001*	1.695	1.288	2.232
	2=secondary and higher	0.130	0.162	0.645	1	0.422	1.139	0.829	1.566
Residence	0=Urban(ref.cat)	-	-	-	-	-	-	-	-
	1=Rular	-1.385	0.136	103.25	1	≤ 0.001*	0.250	0.192	0.327
Age at first sex				12.271	4	0.015			
	0=Not had sex(ref.cat)	-	-	-	-	-	-	-	-
	1=< 18 year	0.155	0.191	0.658	1	0.417	1.167	0.803	1.696
	2=18-24	0.004	0.162	0.001	1	0.980	1.004	0.730	1.381
	3=25-34	-0.460	0.234	3.879	1	0.049	0.631	0.399	0.998
	4=greater than or equal to 35 year	1.495	0.752	3.954	1	0.047	4.458	1.022	19.450
Constant		-4.791	0.542	78.017	1	≤ 0.001*	0.008		

ref.cat=reference category

Likelihood-Ratio Test

It tests the null hypothesis that $\beta=0$, $\beta=(\beta_1, \beta_{2,\ldots,} \beta_k)^T$. The result presented in Table 3 showed a likelihood ratio test statistic G^2 =493.476 which is distributed as chi-square with 22 degrees of freedom. The tabulated value is X^2_{α} (22) = 33.924, since $G^2 > X^2_{0.05}(22)$, we reject the null hypothesis and conclude that the overall model is good.

Table 3: Overall model evaluation using the Likelihood ratio test

	-2Log likelihood	Likelihood ratio(G2)	d.f	X2α (22)	P-value
Null model	4221.337 3727.861	493.476	22	33.924	≤ 0.001*

d.f=degree freedom, α=5%

The result presented in Table 4 shows that the Hosmer-Lemeshow goodness-of-fit test statistic is not significant. Thus, we do not have evidence to reject the null hypothesis, suggesting that the model fitted the data well.

Table 4: Hosmer and Lemeshow Goodness of Fit Test

Chi-square	D.f	Sig
13.194	8	0.105

3.2 Interpretation of Logistic Regression Coefficients

After the assessment of the overall model evaluation and goodness of fit test, statistical tests of individual predictors were conducted to identify the associated risk factors of HIV status among individuals. The statistical significance of individual regression coefficients is tested using the Wald chi-square statistic.

The result displayed in Table 2 revealed that region, HIV Awareness, place of residence, sex, Age, Educational level, and Age at first sex are significantly associated with HIV status. The other risk factors such as wealth index, marital status, and **multiple sex partner** are not associated with of HIV status of individuals. The predicted probability of an individual being HIV positive with no predictor or constant is 0.0082.

The prevalence of HIV infection was significantly associated with geographical regions. The odd of HIV infections in the Gambela region are found to be 13.349 (95% C.I 5.712-31.196) times the odds of HIV infection in the Somalia region. The odd of HIV infections in Addis Abeba is 4.722 times the odd of individuals living in the Somali region. The odd of HIV infections in the Harare region is found to be 1.732(95% C.I 1.039-2.887) times more likely infected by HIV than in the Somali region. Individuals who live in rural areas were 75 % (OR: 0.250) less likely infected with HIV infection than those individuals who lived in urban areas and its effect is statistically significant. Respondents who have no awareness about HIV 2.491(95% C.I 1.007-6.146) Times more infected than respondents have HIV awareness. The sex of an individual is significantly associated with HIV status. The female individual is about 1.938 times more likely to be infected by HIV than male individuals. Age is significantly associated with associated risk factors of HIV infection. The age of respondents is increase with the increasing prevalence of HIV infection. The odds of HIV infection for age group 15-24 and 25-34 is 80.8% (OR: 0.192) and 34.8% (OR: 0.652) respectively less compared to individuals greater than or equal to 35 years (reference category) and its effect is statistically significant. The odd of HIV-infected individuals who are in primary level education is 1.695 times the odds of HIV infection of in no education or preschool level of education and its effect is statistically significant. The odds of being HIV infected for an individual whose age was greater than or equal to 35 years at first sex was 4.458 times the odds of HIV infection of individuals who had no sex.

Results of Multilevel Logistic Regression Analysis

In the multilevel analysis, a two-level structure is used with regions as the second-level units and individuals as the first-level units. This is the analysis of region-wise variation of associated risk factors of HIV status. Individuals were nested in regions with a total of 26692 individuals included in this study.

Test of Heterogeneity

A chi-square test statistic was applied to assess heterogeneity in the proportion of individuals who had HIV infection among regions. The test yield $\chi 2 = 237.960$, d.f=10, P<0.05. Thus, there is evidence of heterogeneity of HIV status of individuals among regions.

Multilevel Logistic Regression Model Comparison

Table 5 Multilevel logistic regression model for associated risk factors of HIV status and their deviance based chi-square test statistics.

	Empty model	Random intercept model	Random coefficient model
-2*log-likelihood	4052.5	3758.6	3730.0
Deviance based chi-square test	168.82	293.9	28.636
P-Value	≤ 0.001*	≤ 0.001*	0.951

*significant at 5% level

The deviance-based chi-square value for the empty model shown in Table 5 is the difference in log-likelihoods between an empty model of single-level logistic regression and an empty model of multilevel logistic regression, which is to be compared with the critical value from the chi-squared distribution with 1 degree of freedom.

Thus, the deviance-based chi-square test in Table 5 shows that among multilevel logistic regression models, the random intercept and fixed slope model fits significantly better than the other multilevel logistic regression models. **Results of Empty Multilevel Logistic Regression Model**

The empty model contains no explanatory variables and it can be considered as a parametric A version of assessing heterogeneity of HIV Status of individuals among regions.

Table 6: Results for Multilevel logistic regression model without explanatory variables				
Fixed part	**Coefficients**	**S.E.**	**Wald**	**p-value**
βo – intercept	-4.318	0.231	-18.64	≤ 0.001*
Random part	Estimate	S.E.	Wald	P-value
Region σu2 = var(uoj)	0.547	0.004	136.75	≤ 0.001*
Rho(p)			0.083	
Deviance based chi-square ≤ 0.001*			168.82	

Significant at 5%.

Table 6 shows the output of the estimates of fixed effects and random effects. From the table, we can see that the estimate of the fixed part of the model is -4.318 with Wald Statistics -18.64 and a p-value of ≤ 0.001. The table also contains the variance estimate of random effects at the regional level, σ^2_u =0.547 which implies that between region variance of HIV infection is 0.547. At the bottom of table 4.6 is the result of the hypothesis. We see that the value of the test statistic is 168.82 with a p-value<0.0001*. Therefore, the null hypothesis is rejected, then the variance of the random factor in the empty model is significant which indicates that there are significant regional differences in the associated risk factors of HIV Status. Based on the significant LR the chi-square value for the no intercept model implies that an empty model for HIV status with random effect is better than an empty model for HIV status without random effect.

Result of Random Intercept and Fixed Slope Multilevel Logistic Regression Model

The random intercept and fixed slope logistic regression model is a multilevel model which has random intercept and fixed coefficient of predictors. As can be seen from Table 6, the analysis of multilevel logistic regression revealed that the HIV status of individuals varied among regions. The value of Var(U_{oj}) is the estimated variance of the intercept in the random intercept and fixed coefficients model. The result displayed that regional variation in the HIV status of an individual is statistically significant. In addition, HIV Awareness, sex, residence, educational level, and Age at first sex were also found to be significant determinants of variation in HIV status of individuals among regions.

Table 7: Results of a random intercept and fixed coefficient logistic regression model Fixed part							
Fixed effect	$\hat{\beta}$	S.E($\hat{\beta}$)	**Wald**	**exp** ($\hat{\beta}$)	**95% CI**		**P-value**
					Lower	**Upper**	
Constant	-5.212	1.173	-4.442	0.005			≤ 0.001*
HIV awareness							
Yes(ref.cat)							
No	0.923	0.461	1.199	2.516	0.019	1.826	0.045*
Sex							
Male(ref.cat)							
Female	0.666	0.111	5.968	1.946	0.449	0.883	≤ 0.001*
Residence							
Urban(ref.cat)							
Rural	-1.306	0.167	-7.794	0.270	-1.633	-0.979	≤ 0.001*
Educational level							
Noeducation,perschool(ref.cat)							
Primary	0.524	0.140	3.730	1.688	0.25	0.798	≤ 0.001*
Secondary and higher	0.130	0.162	0.806	1.138	-0.187	0.447	0.4204
Age at first sex							
Not had sex(ref.cat)							
< 18 years	0.294	0.215	1.366	1.341	-0.127	0.715	0.171
18-24	0.124	0.197	0.630	1.132	-0.262	0.510	0.528

25-34	-0.359	0.257	-1.396	0.698	-0.862	0.144	0.162
> or equal to 35 years	1.579	0.755	2.091	4.850	0.1	3.058	0.0365*
Random part	Estimate	S.E	Wald	$\exp(\hat{\beta})$	95% CI		P-value
Random intercept: σ20= var(uoj)	0.282	0.0032	88.125	1.325	0.276	0.288	≤ 0.001*
Intra-region correlation				0.078			
Deviance based chi-square				293.9			

*significant at 5% level, (ref.cat) = reference category, ICC: Intra-region correlation

In Table 7 the variance component representing variation between regions has decreased from 0.547 in the empty model with a random intercept to 0.282 in the random intercept and fixed slopes multilevel logistic regression model and the significance of it indicates that there is a significant variation between regions in associated risk factors of HIV infection. To correctly interpret the parameter estimates related to predictors in a multilevel model, it is more meaningful to state that the individual estimates increase or decrease the log odds of the outcome [14]. The result displayed in Table 7, indicated that individuals who live in rural areas are less likely to have HIV infection. Individuals who are female and primary educational level, age at first sex greater than or equal to 35 years and individuals who have no HIV awareness more likely to have HIV infection than male individuals, no education, preschool level of education, have no HIV Awareness, respectively.

The intra-region correlation coefficient (rho) is a measure of HIV status within the region. The result presented in Table 7 is 7.8% of total variability HIV status was due to variations within regions.

Discussions

This study aimed to identify the demographic, socio-economic, HIV/AIDS Awareness, and sexual behavior factors related to the HIV status based on Ethiopian Demographic and Health Survey data set.

Residential differences had a significant impact on the prevalence of HIV/AIDS. This finding is in agreement with a study done in sub-Saharan Africa by [15] the results of this study show that for both males and females, the risk of being HIV seropositive was relatively higher among urban residents. The odd of respondents who lived in urban areas were to be HIV positive than the odd of rural respondents. Concerning regional variations, HIV prevalence is highest in the Gambela regions. The odd of HIV infections of individuals in the Gambela region is found to be 13.349 (95% C.I 5.712-31.196) times odd of HIV infections of individuals in the Somalia region. This finding is in agreement with a study done in Ethiopia by [16, 17], study on HIV Prevalence Correlates with High-Risk Sexual Behavior in Ethiopia's Regions. HIV prevalence varies in Ethiopia's eleven regions. Female clients are more likely to be HIV positive than males. This finding is in agreement with a study done in sub-Saharan Africa by [18]. Vulnerability of women to HIV this study may indicate that biological factors are important in explaining the higher HIV prevalence among women than men in Ethiopia. HIV prevalence is higher in Respondents whose Age at first sex is greater than or equal to 35 years. Also, the predicted probability of individuals to be HIV positive with predictor variable age is greater than or equal to 35 years of Age at First sexual intercourse is 0.988. This finding is in agreement with a study done in Ethiopia by [19] that the odds of being HIV infected for an individual aged ≥ 35 years have 1.327 more than those in the age group $\leq 18\ years$ $(95\%\ C.I = 0.980 - 1.798)$ and is statistically significant at 0.1 level of significance. This was probably due to unsafe sexual practices for the aged individual.

HIV prevalence is higher in educated individuals than Individuals who had no education or preschool level of education. The odd of HIV infections of individuals who are in primary level education is 1.695 times than the odd of HIV infection of individuals with no education or preschool level of education. The predicted probability of individuals being HIV positive with a primary level of education is 0.844. The study conducted by Jukes et al. [20] Using Cross-sectional design showed that educational attainment is associated with a higher risk of HIV infection Existing literature suggests mixed patterns in the association of HIV infection and education status. those with the highest levels of education were found to be more likely to be infected with HIV than those at the lower end of the education spectrum for the reasons that the more educated were wealthier, more mobile and had broader networks of sexual partners [21].

HIV/AIDS awareness is significantly associated with HIV infection. The odd of respondents who have no HIV awareness is 2.491(95% C.I 1.007-6.146) Times more infected than the odd of respondents who have HIV awareness. This finding is in agreement with a study done in sub-Saharan Africa by [15] on "A cross-national analysis of factors associated with HIV infection in sub-Saharan Africa: evidence from the DHS. In all sub-Saharan countries where there is a significant association between HIV infection

and HIV/AIDS awareness, the general pattern suggests a higher HIV prevalence among those with higher HIV/AIDS awareness, especially among females.

CONCLUSION

HIV testing is a critical first step toward HIV prevention, treatment, care, and support. The goal of this research was to investigate demographic, socioeconomic, HIV/AIDS awareness, and sexual behavior characteristics as drivers of HIV status in Ethiopia, as well as to examine regional heterogeneity in HIV status.

Despite the fact that the frequency of HIV infection varies by area. People who lived in the country's cities were at a higher risk of contracting HIV. Females are more likely than males to get infected with HIV. Many additional research have found a similar trend of female sensitivity to HIV/AIDS infection in Ethiopia. HIV affects those with a greater level of education more than people who have just attended preschool. Individuals whose age at first sex is higher than or equal to 35 years are the most affected age group by HIV/AIDS in terms of age at first sex. The prevalence of HIV is inversely proportional to HIV awareness. Individuals who are HIV-aware are more likely to be infected than those who are HIV-unaware. According to the findings, the government should provide more support and attention to areas with high incidence of HIV/AIDS infection. More research into demographics, socio-cultural habits, and other associated issues is needed.

Abbreviation

AIDS: Acquired Immune Deficiency Syndrome; AGYW: Adolescent Girls and Young Women CSA: Central Statistical Agency; DHS: Demographic and Health Survey; EA: Enumeration Area; EDHS: Ethiopia Demographic and Health Survey; EPHI: Ethiopia Public Health Institute; HIV: Human Immune Deficiency Virus; LR: Likelihood-Ratio; OSSA: Organization for Social Services for AIDS; RC: Relative Concentration; SNNPR: South Nation Nationalities and People Regional State; WHO: World Health Organization

Acknowledgments

The authors are glad to thank the MEASURE DHS Project for free access to the Ethiopia Demographic and Health Survey 2016 data.

Authors' Contributions

All SS, SB, and AT carried out the data extraction, designed the study, performed analysis and interpretation, and draft the manuscript. also, all authors designed the study, participated in the analysis and interpretation of the drafted manuscript. All authors read and approved the final manuscript.

Funding

Not applicable

Availability of data and materials

The data for this study were sourced from Demographic and Health Surveys (DHS) and are available at http://www.dhsprogram.com/data/available-datasets.cfm.

Declarations

Ethics approval and consent to participate

Written consent was obtained from the Measure DHS International Program which authorized the data sets. The data used in this study were publicly available, aggregated secondary data with not having any personal identifying information that can be linked to particular individuals, communities, or study participants. Confidentiality of data is maintained anonymously.

Consent for Publication

This manuscript has not been published elsewhere and is not under consideration by any other journal. Both authors approved the final manuscript and agreed with its submission. We agreed about authorship and the order of authors for this manuscript.

Competing interests: The authors declare there was no competing interest.

Reference

1. Chandrasekaran P, Dallabetta G, Loo V, et al. Containing HIV/AIDS in India: the unfinished agenda. The Lancet infectious diseases. 2006;6(8):508-21.
2. Appiah A, Gates HL. Encyclopedia of Africa: Oxford University Press on Demand, 2010.
3. UNAIDS. UNAIDS 2016–2021 Strategy: On the fast-track to end AIDS. UNAIDS Geneva; 2015.
4. EDHS E. demographic and health survey 2016: key indicators report. The DHS Program ICF. 2016.
5. UNAIDS U. The gap report. Geneva, Switzerland. 2014.
6. Hosmer DW, Lemeshow S, Sturdivant RX. The multiple logistic regression model. Applied logistic regression. 1989;1:25-37.
7. Guo G, Zhao H. Multilevel modeling for binary data. Annual review of sociology. 2000;26(1):441-62. https://doi.org/10.1146/annurev.soc.26.1.441.
8. Snijders TA, Bosker RJ. An introduction to basic and advanced multilevel modeling. Sage, London WONG, GY, y MASON, WM (1985): The Hierarchical Logistic Regression Model for Multilevel Analysis, Journal of the American Statistical Association. 1999;80(5):13-524. https://doi.org/10.1080/01621459.1985.10478148.
9. Sommet N, Morselli D. Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using stata, R, Mplus, and SPSS. International Review of Social Psychology. 2017;30:203-18. https://doi.org/10.5334/irsp.90.
10. Goldstein H. Multilevel modelling of educational data. Methodology and epistemology of multilevel analysis: Springer, 2003. p. 25-42. https://doi.org/10.1007/978-1-4020-4675-9_2.
11. Goldstein H, Rasbash J. Improved approximations for multilevel models with binary responses. Journal of the Royal Statistical Society: Series A (Statistics in Society). 1996;159(3):505-13. https://doi.org/10.2307/2983328.

12. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. Journal of the American statistical Association. 1993;88(421):9-25. https://doi.org/10.1080/01621459.1993.10594284.
13. Hosmer DW, Lemeshow S. Applied Logistic Regression. John Wiley & Sons. New York. 2000. https://doi.org/10.1002/0471722146.
14. Badlam JB, Bull TM. Steps forward in the treatment of pulmonary arterial hypertension: latest developments and clinical opportunities. Therapeutic advances in chronic disease. 2017;8(2-3):47-64. https://doi.org/10.1177/2040622317693218.
15. Magadi M, Desta M. A multilevel analysis of the determinants and cross-national variations of HIV seropositivity in sub-Saharan Africa: evidence from the DHS. Health & place. 2011;17(5):1067-83. https://doi.org/10.1016/j.healthplace.2011.06.004.
16. Kenyon CR, Tsoumanis A, Schwartz IS. HIV Prevalence Correlates with high-risk sexual behavior in Ethiopia's Regions. PloS one. 2015;10(10):e0140835. https://doi.org/10.1371/journal.pone.0140835.
17. Paul R, Mukkadan JK. Modulation of blood glucose, oxidative stress, and anxiety level by controlled vestibular stimulation in prediabetes. J Nat Sci Biol Med. 2020;11:111-7.
18. Hegdahl HK, Fylkesnes KM, Sandøy IF. Sex differences in HIV prevalence persist over time: evidence from 18 countries in sub-Saharan Africa. PloS one. 2016;11(2):e0148502. https://doi.org/10.1371/journal.pone.0148502.
19. Danna DD, Badacho AS, Taye A. Statistical Analysis on Correlates of HIV Testing Outcomes Based on Ethiopia Demographic and Health Survey (EDHS) 2011 Data. Science. 2018;6(5):120-6. https://doi.org/10.11648/j.sjph.20180605.11.
20. Jukes M, Simmons S, Bundy D. Education and vulnerability: the role of schools in protecting young women and girls from HIV in southern Africa. Aids. 2008;22:S41-S56. https://doi.org/10.1097/01.aids.0000341776.71253.04.
21. Hargreaves JR, Bonell CP, Boler T, et al. Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. Aids. 2008;22(3):403-14. https://doi.org/10.1097/QAD.0b013e3282f2aac3.