

Application of Binary Logistic Regression in Assessing Risk Factors Affecting the Prevalence of Toxoplasmosis

Mutangi Kudakwashe^{1,*}, Kasim Mohammed Yesuf²

¹Department of Mathematics, Harare Institute of Technology, P O box BE277, Belvedere, Harare, Zimbabwe

²Kasim Mohammed Yesuf, Department of Statistics, University of Gondar, P O Box 196, Gondar, Ethiopia

*Corresponding author: kmutangi@gmail.com

Received August 14, 2014; Revised October 14, 2014; Accepted November 11, 2014

Abstract Toxoplasmosis is a parasitic disease caused by the protozoan parasite *Toxoplasma Gondii* (*T.gondii*). The parasite infects warm-blooded animals among them humans especially those whose immunity has been compromised. The transmission mode of the parasite vary from living in unhygienic conditions, contact with cat faeces to contact with raw meat or the practice of raw meat eating, such as commonly practiced in Ethiopia. Binary logistic regression was used to determine the risk factors affecting the prevalence of toxoplasmosis in HIV/AIDS patients. Significant risk factors were detected using the Wald and the likelihood ratio tests. The model selected was then subjected to diagnostic checks to assess its fitness using the Hosmer and Lemeshow test as well as the Pearson, and deviance goodness of fit tests. The results showed that patients living under unhygienic conditions, aged patients, illiterate and less educated patients were mostly affected by toxoplasmosis. There was more prevalence in urban areas than in rural areas possibly due to the high density of people in urban areas.

Keywords: logistic regression, toxoplasmosis, risk factors, prevalence, diagnostic tests, HIV/AIDS

Cite This Article: Mutangi Kudakwashe, and Kasim Mohammed Yesuf, "Application of Binary Logistic Regression in Assessing Risk Factors Affecting the Prevalence of Toxoplasmosis." *American Journal of Applied Mathematics and Statistics*, vol. 2, no. 6 (2014): 357-363. doi: 10.12691/ajams-2-6-1.

1. Introduction

T. gondii is a protozoan parasite which is an endemic world-wide organism [16]. It has been estimated that about one third of the human population is infected with this parasite [6]. Toxoplasmosis is arguably the most opportunistic disease in HIV infected persons. There is a wide geographic variation in the prevalence of latent toxoplasma infection [21].

Toxoplasmosis if not properly managed has the capacity to cause morbidity and mortality in immune-compromised patients. This disease generally results from reactivation of latent infection [12]. Although *T.gondii* can cause disease with multiple sites of infection, one organ may be preferentially infected. About 4% of all cases of pneumonia in AIDS patients can be attributed to pulmonary toxoplasmosis and is the second or third most frequent form of toxoplasmosis after toxoplasmic encephalitis [12]. Toxoplasmosis is normally asymptomatic in healthy individuals but can cause abortion in pregnant patients. The overall risk for maternal-fetal transmission in HIV negative women who acquire primary toxoplasma infection during pregnancy is 29% [10]. Among the congenital infections, it is estimated that 10-13% of the babies will have visual handicaps.

Acute cerebral toxoplasmosis is the most common cause of focal neurologic disorder in AIDS patients. If not detected and treated promptly, cerebral toxoplasmosis

may cause significant morbidity and mortality. All HIV-infected persons should be made aware of the non-pharmacologic and medical prophylaxis for *T. Gondii* infection and sero-positive patients should receive either primary or secondary prophylaxis for toxoplasmosis [11].

The sero-prevalence of anti-toxoplasma IgG antibody among HIV+ and HIV negative participants was determined by the study in Addis Ababa, Ethiopia. The overall prevalence rate of latent *T.gondii* infection was found to be 90%. The high sero-prevalence of latent toxoplasma infection among the study population seems reasonable as raw or insufficiently cooked meat is consumed in various cultural food outlets. Cats are also abundant and the climatic conditions favour the survival of the parasite [16]. In Ethiopia, the prevalence of IgG antibodies to toxoplasma gondii has been determined by an enzyme-linked immunosorbent assay (ELISA). One thousand and sixteen sera collected in six different geographical regions were analysed. Antibody titres >15 IU/ml were detected in 74.4% of the specimens, titres exceeding 200 IU/ml in one third of the ELISA-positive sera [8].

The association between socio-economic characteristics of a society and the provision of health care has become a matter of interest among several authors. Inequalities have been reported with particular contrasts between rural and urban areas [19]. Levels of urban health are sometimes viewed as being worse than in rural areas, but contrary evidence do also exist [4]. Health is believed to be influenced by ecological, social environment, genetics and

individual characteristics. Urban and rural differences in mortality are determined by many factors such as differences in lifestyle, ecological situations and access to health services, unequal distribution of incomes and resources [17]. Rural people are less likely to protect themselves against HIV and if they become ill, they are less likely to receive adequate care. Poverty widespread in rural areas leads to poor nutrition and poor health which makes a person more vulnerable to HIV related death [17].

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. Sometimes the outcome variable is discrete with two or more possible values. Binary logistic regression has become the standard method of analysis in situations where the outcome variable is discrete. Early uses of binary logistic regression were in biomedical sciences but the past twenty years has also seen much use in social sciences and marketing [1].

Logistic regression is an increasingly popular statistical technique used to model the probability of discrete (binary or multinomial) outcomes. When properly applied, logistic regression analyses yield very powerful insights into which variables are more or less likely to predict an event outcome in a population of interest. Modelling a binary response variable using normal linear regression introduces substantial bias into parameter estimates.

1.1. Statement of the Problem

In 2008, an estimated 1.7 million deaths (2%) in Ethiopia were due to HIV/AIDS related illness in the 15-49 age groups which is also the productive age group. In 2010 the figure rose to 1.9 million deaths in the same age group [5]. Toxoplasma infection in humans may be acquired by eating undercooked infected meat containing Toxoplasma cysts, ingestion of the oocysts from faecally contaminated hands or food, organ transplantation or blood transfusion, transplacental transmission and accidental inoculation of tachzoites. In Ethiopia, there is a practice of raw meat eating which leaves the population exposed to toxoplasmosis.

[2] focused on the prevention and factors that increase the chances of contracting toxoplasmosis, giving less attention to improving the livelihoods of HIV+ patients on ART. Several other researchers didn't focus on assessing the factors that influence the survival status of HIV+ persons on ART who might be exposed to toxoplasmosis. Earlier researchers also tended to concentrate on factors affecting toxoplasmosis without taking their time to employ rigorous statistical analysis coupled with adequate model diagnostic checks. A binary logistic regression model was therefore used for the investigation of risk

factors that affect toxoplasmosis prevalence in HIV/AIDS patients. This is the motivation for this research among other factors already mentioned.

1.2. Significance of the Study

This study will result in the following benefits to the society (health professionals, academics, researchers, general public)

- It will give a statistical method useful in determining the risk factors affecting the prevalence of toxoplasmosis.
- It will provide the risk factors themselves.
- Promote effective planning and policy useful on prevention, control and intervention strategies of toxoplasmosis.
- Results of this investigation will be used as a basis for further studies on toxoplasmosis prevalence in populations with the same characteristics as the one involved in this study.

1.3. Objectives of the Study

- To find out the risk factors that affects the prevalence of toxoplasmosis.
- To identify the best binary logistic model for the data using the logit, probit (normit) link functions and the complementary log log (cloglog) link function.
- To assess the model for adequacy using existing diagnostic tools.

2. Materials and Methods

2.1. Data Collection

The data was obtained from Jimma University Research and publications office. The study included 120 subjects from all HIV/AIDS patients who came to Mettu Karl hospital laboratory for CD4 counts and ART monitoring during the study period. Mettu town is the capital of Illu Ababor zone about 600 km south west of Addis Ababa. The data was collected by assessing the patient's socio-demographic and nutritional conditions. A questionnaire was used. The physician also collected the relevant clinical data of manifested toxoplasmosis before sending the patient for sample collection for the documentation of the clinical aspect of the study. A cross-sectional study with systematic random sampling was used. Direct Agglutination Test was used to detect toxoplasma antibodies in the subject's serum.

2.2. Variables Included in the Study

Table 1. Independent variables involved in the study

No	Variable	Code	No	Variable	Code
1	Sex of patient	Sex	9	Tap water presence	Tap water
2	Age of patient	Age	10	Compound latrine	Latrine
3	Place of residence	Adr	11	House toilet	Hsetoilet
4	Education	Educ	12	Cat presence	Cat
5	Occupation	Occ	13	Raw meat eating	Rawmeat
6	Rooms in house	Rooms	14	Hand washing	Hand
7	Marital status	Mar	15	Fruit cleaning	Fruit
8	Family size	persons	16	Feeling fever	Fever

2.2.1. Dependent Variable

The dependent variable was the toxo status of an HIV/AIDS patient (positive=1, negative=0), it is denoted TOXO.

2.2.2. Independent Variables

Sixteen variables were considered, variables assumed to influence the prediction of positive/negative toxo status.

2.3. Methodology

Since our dependent variable is dichotomous with several independent variables, logistic regression is preferred from multiple regression and discriminant analysis. Logistic regression apart from being mathematically flexible and easy to use, it results in a biologically meaningful interpretation [9]. In this section we discuss the logistic regression with the logit, probit and the complementary log log link functions.

2.3.1. Binary Logistic Regression Model

The binary logistic regression is discussed under three link functions (2.3.2, 2.3.3 and 2.3.4) as stated above

2.3.2. Logit Model

The logit model is described as follows:

$$\text{Logit}(P(Y = 1)) = \beta_0 + \beta_{1X_{i1}} + \sum_{k=1}^3 \beta_{2k} X_{2ki} + \beta_3 X_3 + \dots + \sum_{k=1}^4 \beta_{16k} X_{16ki} \quad (1)$$

Where X_j is the j^{th} risk factor, K is the k^{th} category of the factor, β_0 is the intercept of the logit model, β_j is the estimated coefficient for each risk factor j for the logit model and $p(Y=1)$ is the probability that the i th patient will develop toxoplasmosis.

2.3.3. The Probit (Normit) Model

Using the probit link function, the binary logistic model is described as the inverse of the normal distribution function

$$\text{Probit}(P(Y = 1)) = \Phi^{-1}(P(Y_i = 1)) \quad (2)$$

2.3.4. Complementary Log Log (cloglog) Model

The cloglog link function results in a model described as follows:

$$\text{Log}(-\log(1 - P(Y = 1))) = \beta_0 + \beta_{1X_{i1}} + \sum_{k=1}^3 \beta_{2k} X_{2ki} + \beta_3 X_3 + \dots + \sum_{k=1}^4 \beta_{16k} X_{16ki} \quad (3)$$

2.4. Model Building Strategies (Variable Selection)

The procedure used was to fit an intercept only model (null model) and then fit a model with more parameters (fitted model). The following hypothesis was then tested for significance using the likelihood ratio test: $\chi^2_{LR} = -2\log(L_0 / L_1)$, where L_0 is the maximised value

for the null model and L_1 is the maximised value for the fitted model.

$H_0 : \beta_j = 0$ for all j versus $H_1 : \beta_j \neq 0$ for at least one j .

Rejection of the null hypothesis would mean that at least one of the parameters is different from zero. The Wald test was used to test the statistical significance of each of the parameters β_j once the null hypothesis is rejected.

One problem with any bivariate approach is that it ignores the possibility that a collection of variables, each of which is weakly associated with the outcome, can become an important predictor of the outcome. In using the package R, the logistic regression is a stepwise procedure which begins by selecting the strongest candidate predictor and then testing additional candidate risk factors (predictors) one at a time for inclusion in the model. This procedure offers several methods for stepwise selection of the best risk factors to include in the model based on the Akaike information Criteria (AIC). The Logit, probit and Cloglog link functions were employed.

2.5. Interpretation of Parameters

The model is monotone depending on the sign of β . $P(Y=1)$ is increasing if β is positive and decreasing otherwise as X increases by one unit. For the logit model, the odds are an exponential function of X . Thus

$$\frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \exp[\beta_0 + \beta_{1X_{i1}} + \sum_{k=1}^3 \beta_{2k} X_{2ki} + \beta_3 X_3 + \dots + \sum_{k=1}^4 \beta_{16k} X_{16ki}]$$

Is the odds of developing toxoplasmosis for the i^{th} patient. For every unit increase in X the odds increases multiplicably by e^{β_j} . The probit and the cloglog models do not provide an estimate of the odds ratio.

2.6. Assessing the Fit of the Model

In assessing the quality of the fitted models, Pearson's chi-square statistic and the likelihood ratio statistics (G^2) which are based on the comparison of the fitted and observed counts were used. If the chi-square is significant, the variable is considered to be significant in the model. The Hosmer-Lemeshow test was also employed. The Likelihood ratio test statistic $\chi^2_{LR} = -2\log(L_0 / L_1) = -2(\log L_0 - \log L_1)$ was also used where L_0 is the maximised value for the null model and L_1 is the maximised value of the likelihood function for the fitted model.

3. Results

3.1. Summary Statistics

3.1.1. Socio-economic Factors

About 72.5% of patients with no access to tap water tested positive to toxoplasmosis with 27.5% testing negative. 87.5% of farmers tested positive to toxoplasmosis while only 25% of merchants had the same results. Patients living in 1-2 rooms were very much exposed with 69.2% of them testing positive.

3.1.2. Demographic and Health Factors

Out of 120 patients considered in the study, 68 (59.6%) tested positive for toxoplasmosis with 76.5% of males testing positive compared to 52.5% females. The 18-23 year age group was more exposed than other age-groups with 69.2% of them testing positive. Illiterate and semi-literate people showed high exposure with 70.6% and 75% respectively testing positive for toxoplasmosis. Family sizes of 3 and below showed high proportions with 61.1% testing positive. 63.2% of patients with high fever also tested positive for toxoplasmosis. Living under unhygienic conditions exposed patients to toxoplasmosis with 81.2%

of patients without a compound latrine testing positive. 62.5% of patients who lived with cats also tested positive.

3.1.3. Risk Behaviour Factors

Eating raw meat had a higher exposure to toxoplasmosis with 67.9% of raw meat eating patients testing positive. Patients who had the habit of washing their hands and fruit before eating had a low risk with 48.8% and 46.9% respectively testing positive.

3.2. Results on Tests of Association between the Independent Variables and the Dependent Variable (TOXO)

Each of the independent variables was tested for association with the dependent variable (toxoplasmosis) taking one variable at a time. The Pearson Chi-square (X^2) and the likelihood ratio (G^2) tests of association were used. The results are in Table 2 below:

Table 2. Tests of association between toxoplasmosis and each of the independent variables

No	Variable	Chi-square test		Likelihood ratio test	
		Value	p-value	Value	p-value
1	Sex of patient (SEX)	5.696	0.017*	5.961	0.015*
2	Age of patient (AGE)	2.421	0.490	2.438	0.487
3	Place of residence (ADR)	0.009	0.923	0.009	0.923
4	Occupation (OCC)	11.247	0.024*	12.096	0.017*
5	Marital Status (MAR)	5.331	0.070	5.909	0.052
6	Education (EDUC)	12.461	0.006*	12.463	0.006*
7	Number of family members (PERSONS)	0.174	0.677	0.173	0.677
8	Number of rooms in house (ROOMS)	9.422	0.002*	9.362	0.002*
9	Presence of tap water (TAPWATER)	18.405	0.000*	18.464	0.000*
10	Compound latrine (CPDLATRINE)	8.624	0.003*	8.549	0.003*
11	House toilet (HOUSETOILET)	6.15	0.013*	6.053	0.014*
12	Cat presence (CAT)	12.44	0.000*	12.485	0.000*
13	Raw meat eating (RAWMEAT)	3.081	0.079	3.100	0.078
14	Habit of washing hands (HAND)	14.337	0.000*	16.025	0.000*
15	Habit of washing fruits (FRUIT)	9.893	0.002*	10.185	0.001*
16	Feeling fever (FEVER)	0.292	0.589	0.293	0.588

*significant (p-value<0.05)

The results above indicate that the following variables influence toxoplasmosis in a significant way: SEX, OCC, EDUC, ROOMS, TAPWATER, CPDLATRINE, HOUSETOILET, CAT, HAND and the practice of washing fruits.

3.3. Results from Multiple Logistic Regression Analysis

3.3.1. Variables Identified by the Stepwise Selection Using the Logit Link Function

Table 3. Variables identified by the stepwise procedure using the logit link function

Var	Deviance	AIC	Var	Deviance	AIC
AGE	63.627	91.627	HAND	76.060	108.060
FEVER	61.367	93.367	EDUC	84.701	112.701
FRUIT	62.132	94.132	MAR	87.574	117.574
ROOMS	62.642	94.642	TAPWATER	97.903	129.903
HSETOILET	64.895	96.895	ADR	72.812	104.812
PERSONS	65.049	97.049	None	57.419	91.419

3.3.2. Variables identified by Stepwise Selection Using the Probit Link Function

Table 4. Variables identified by stepwise procedure using the Probit link function

Var	Deviance	AIC	Var	Deviance	AIC
None	62.724	90.724	ADR	73.081	99.081
FEVER	65.449	91.449	HAND	79.434	105.434
FRUIT	65.589	91.589	EDUC	87.118	109.118
ROOMS	66.645	92.645	MAR	92.936	116.936
PERSONS	68.826	94.826	TAPWATER	104.656	130.656
HSETOILET	72.256	98.256			

Table 5. Variables identified by stepwise selection using the cloglog link function

Var	Deviance	AIC	Var	deviance	AIC
None	55.438	89.438	PERSONS	63.865	95.865
AGE	61.673	89.673	ADR	70.665	102.665
FEVER	58.221	90.221	HAND	78.632	110.632
FRUIT	59.104	91.104	EDUC	83.436	111.436
ROOMS	61.345	93.345	MAR	89.389	119.389
HSETOILET	62.419	94.419	TAPWATER	95.774	127.774

3.3.3. Parameter Estimates

Table 6. The estimated coefficients for the covariates in the logit and the probit models

Logit link function				Probit link function			
covariate	$\hat{\beta}$	Wald	Sig	Covariate	$\hat{\beta}$	Wald	sig
intercept	-12.11369	-3.328	0.000874*	Intercept	-6.6651	-3.515	0.00044*
Age24-9	-1.65338	-1.456	0.145451	Adr_urban	3.0078	2.525	0.01156*
Age42+	-3.20999	-2.162	0.030588*	Educ1-4	0.3612	0.492	0.62285
Adr_urban	6.73501	2.739	0.006163*	Educ5-8	-1.5851	-2.162	0.03065*
Educ1-4	-0.07191	-0.049	0.961264	Educ9+	2.1844	2.822	0.00477*
Educ5-8	-2.60734	-1.624	0.104433	Mar_mard	-4.5691	-3.903	0.00000*
Educ9+	4.68080	2.905	0.003673*	Mar_other	0.8535	0.79	0.42969
Mar_mard	-7.00668	-3.269	0.001077*	Rooms3+	1.5748	2.008	0.04463*
Mar_other	3.36102	1.371	0.170220	Persons4+	1.3951	2.292	0.02189*
Rooms3+	3.01428	2.186	0.028786*	Tapwatern	4.6559	4.155	0.0000*
Persons4+	2.80996	2.524	0.011616*	Hsetoiletn	2.0389	2.806	0.00502*
Tapwatern	8.01841	3.610	0.000306*	HandS	4.2171	3.54	0.0004*
Hsetoiletn	3.22313	2.564	0.010348*	FruitS	2.0548	1.634	0.10224
HandS	8.01841	3.283	0.001025*	Fevern	-0.7954	-1.582	0.11374
FruitS	4.52459	1.869	0.061579				
Fevern	-1.80995	-1.829	0.067357				

*implies significant result. HandS=sometimes washing hands with FruitS having similar interpretation. Tapwatern= no tapwater, similar interpretations for fevern.

All the significant parameters were included in their respective models and then checked for adequacy. The AIC for the logit model was 91.419 while that for the probit model was 90.724.

Table 7. Parameter estimates using the cloglog link function

Covariates	$\hat{\beta}$	Wald	Sig	Covariates	$\hat{\beta}$	Wald	Sig
intercept	-9.4016	-3.562	0.00368*	Mar_other	2.3339	1.624	0.10429
Age24-29	-1.3556	-1.863	0.06241	Rooms3+	2.2202	2.614	0.00895*
Age30-35	-0.7547	-1.007	0.31372	Persons4+	2.2998	2.887	0.00389*
Age42+	-2.4572	-2.304	0.02121*	Tapwatern	5.6634	3.582	0.00034*
Adr_urban	5.1724	2.752	0.00593*	Hsetoiletn	2.3647	2.696	0.00702*
Educ1-4	-0.7786	-0.844	0.39859	HandS	6.0985	3.852	0.00011*
Educ5-8	-1.8423	-1.570	0.11648	FruitS	2.7865	1.601	0.10936
Educ9+	3.2322	2.966	0.00302*	Fevern	-1.1207	-1.711	0.08715
Mar_mard	-4.9009	-3.233	0.00122*				

The AIC for the cloglog is 89.438.

3.4. Final Selected Model

In selecting the best model, the selection criteria involved choosing the distribution and link function as well as which covariates to include in the model. The model with the smallest AIC was selected as the best model.

Table 8. AIC values for the three models used in this study

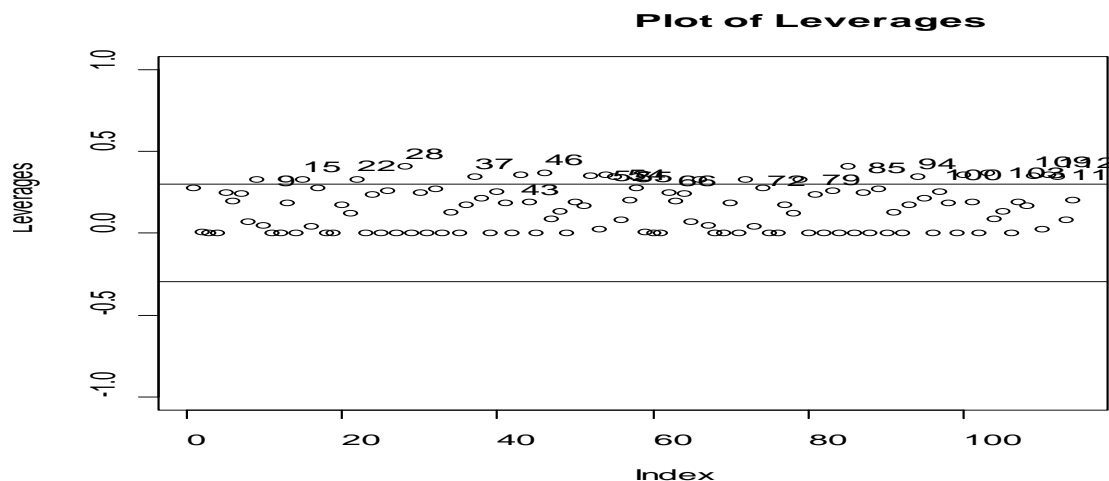
Model	Loglikelihood	No of parameters	AIC
Logit	-28.70936	17	91.419
Probit	-31.36183	14	90.724
Cloglog	-27.71908	17	89.438

The cloglog model was selected since it has the smallest AIC.

3.5. Model Adequacy Checking

A goodness of fit test (G) was used to assess the overall fitness of the selected cloglog model. This is the chi-square difference between the null model (with intercept only) and the model containing one or more predictors.

For our selected cloglog model, $\chi^2_{LR} = 98.33$ with $df=16$ and $p=0.000$ implying a significant increase in the likelihood thereby implying a good fit of the model. The difference between the deviance of the null (M_1) and full (M_2) models yields $\chi^2_{LR} = D_{M_1} - D_{M_2} \sim \chi^2_{(p-1)}$. The deviance for the cloglog model was $D_{M_2} = 55.438$ with ($df=97$) and $D_{M_1} = 153.77$ with ($df=113$), giving $\chi^2_{LR} = 153.77 - 55.438 = 98.327$ with ($df=113-97=16$) as already given. In the Pearson and Deviance tests, the larger the p-value the better fit to the data. The Pearson and deviance test values were respectively 67.78558 and 55.438 with their corresponding p-values as 0.98944 and 0.99977. This indicates no evidence to suggest that data did not come from a population that follows a logistic regression model. There was also no evidence of over-dispersion as the estimated dispersion parameter was 0.6988.



The plot of leverages shows points with high leverages (greater than $(\frac{2p}{n})$). Points 9, 15, 22 etc have high leverages.

3.6. Discussion of Results

This study showed an overall 59.6% seroprevalence of toxoplasmosis for the one hundred and twenty patients at Mettu Karl hospital included in the study. This is lower than the prevalence in Jimma town (83.6%) [20], 81.4% in central Ethiopia [7] but higher than that reported in Nigeria in which 242 HIV positive patients were tested and an overall 41.3% seroprevalence was reported [15].

Women were less affected with 52.5% seroprevalence in sharp contrast with other studies; Seropositivity was higher in females (42.8%) than males (39.2%) in a study in Nigeria [15]. Toxoplasma seropositivity was not associated with age, sex, Art status and CD4 count [14]. In a study involving 330 sera in Addis, toxoplasma infection was higher in males (93.6%) than females (86.6%) supporting the findings in this study [16].

87.5% of the farmers were seropositive possibly due to their constant contact with contaminated soil and failing to wash their hands before food consumption, this is higher than the 73.6% reported by [20]. About 72.5% of the patients with no access to tape water tested positive for toxoplasmosis compared to 75.58% who tested positive at Douala hospital in Cameroon [13].

One of the risk factors was cat presence. 62.5% of those who lived with cats tested positive for toxoplasmosis, a figure quite lower than the 72.6% reported by [13] but higher than reported in [18] in a study involving Thai pregnant women. Other similar results were reported by [7,15,20]. The practice of raw meat eating is quite common in Ethiopia. 67.9% seroprevalence for raw meat eaters from this study compares well with 73.3% seroprevalence for meat eaters in the Douala study of Cameroon [13]. There was however no significant association between toxoplasmosis and the habit of raw meat eating in a study by [20].

A cloglog link function was selected as the best predictive model in this study. Although use of logistic regression is mentioned in other studies [7,20], none of the authors to the best of my knowledge compared the three link functions (probit, logit & cloglog) in order to come up

with a suitable model to predict the risk factors for toxoplasmosis.

4. Conclusions

The bivariate logistic regression with a complementary log log (cloglog) link function was the best model for the Toxoplasma data from Mettu Karl hospital in South West Ethiopia. The risk factors which significantly affect Toxoplasma infection of HIV/AIDS patients are age, patient's place of residence, educational level, marital status, number of family members in the house, number of rooms in the house, tape water, presence or absence of house toilet and the practice of washing hands. The other factors like gender, one's occupation, cat presence and the practice of raw meat eating were insignificant. Older and less educated patients were mostly affected by toxoplasma gondii. The situation was the same for the married patients and those with no house toilets. The absence of house toilets and tape water means less hygienic conditions at home thus exposing the inhabitants to T. Gondii infection. Recommendations on the prevention of human toxoplasmosis have been suggested by [3] and new research could perhaps focus on finding means to reduce the spread of T.gondii to humans. Prevalence rates of infection of new born babies whose mothers were infected with T.gondii are unknown in Ethiopia to the best of my knowledge and that needs further research.

Acknowledgements

The authors of this article would like to the thank Jimma University research and publications office for making available the data used in this research. We also thank Harare Institute of Technology for the use of their computers.

Statement of Competing Interests

The authors of this study have no competing interests.

List of Abbreviations

Abbreviation	Explanation
ART	Anti-retroviral Treatment
CLOGLOG	Complementary log log link function
FEVERN	No Fever
FRUITS	Sometimes washing Fruits
HANDS	Sometimes washing hands
HSETOILET	House Toilet
HSETOILETN	No House Toilet
IgG	Immunoglobulin G
IgM	Immunoglobulin M
MAR_Mard	Marital status is "Married"
TAPEWATERN	No tap water
TOXO	Toxoplasmosis

References

- [1] Agresti, A, *Categorical data analysis*, Gainesville, Florida, 2002
- [2] Alemtsehai, A. And Eshetu, W, Association of HIV infection with some selected factors and modelling the chance of contracting HIV: The case of Hawassa and its surrounding. *Journal of Ethiopian Statistical Association*, 15. 27-48. 2006
- [3] Bresciani, K.D.S, Galvao, A.L.B, de Vsconcellos, A.L, Soares, J.A, de Matos, L.V.S, Pierucci, J.C, Neto, L.S, Rodrigues, T.O, Navarro, I.T, Gomes, J.F and da Costa, A.J, Relevant aspects of human toxoplasmosis. *Research Journal of Infectious diseases*. 2013. [online]. Available: <http://www.hoajonline.com/journals/pdf/2052-5958-1-7.pdf> more details.
- [4] Buve, A., Carael, M. and Hayes, R.J, Multi-center study on factors determining differences in rate of spread of HIV in sub-saharan Africa: Methods and prevalence of HIV infection. *AIDS* 15 (4). 5-14. 2001.
- [5] Ethiopia Demographic Health Survey (EDHS). Central statistical agency, Addis Ababa, Ethiopia. September 2006.
- [6] Fong, M.Y, Wong, K.T, Rohela, M, Tan, L.H, Adeeba, K., Lee, Y.Y, and Lau, Y.L, unusual manifestation of cutaneous toxoplasmosis in an HIV-positive patient. *Tropical biomedicine*, 27 (3). 447-450. 2010.
- [7] Gebremedhin, E.Z, Abebe, A.H, Tessema, T.S, Tullu, K.D, Medhin, G, Vitale, M, Di Marco, V, Cox, E and Dorny, P, Sero-epidemiology of toxoplasma gondii infection in women of child bearing age in central Ethiopia. *BMC infectious diseases*, 13: 101, February 2013. [online]. Available: <http://www.biomedcentral.com/1471-2334/13/101>. [Accessed July. 25, 2014].
- [8] Guebre-Xabier, M, Nurilign, A, Gebre-Hiwot, A, Hailu, A, sissay, Y, Getachew, E, Frommel, D, Sero-epidemiological survey of toxoplasma gondii infection in Ethiopia. *Ethiopian Medical Journal*, 31 (3). 201-208. 1993.
- [9] Hosmer, D.W, and Lemeshow, S, *Applied logistic regression 2nd edition*, John wiley & Sons Inc, New York, 2000.
- [10] Jara, M, Hsu, H.W, Eaton, R.B, Demaria, A.Jr, Epidemiology of congenital toxoplasmosis identified by population based newborn screening in Massachusetts. *Paediatr infect Disease Journal*, 20 (12). 1132-5. 2001.
- [11] Jayawardena, S, Singh, S, Burzyantseya, O and Clarke, H, Cerebral toxoplasmosis in adult patients with HIV infection. *Hospital Physician*, 44 (7). 17-24. 2008.
- [12] Luft, B.J, and Remington, J.S, AIDS commentary. Toxoplasmosis encephalitis. *Journal of infectious Diseases*, 157 (1). 1-6. 1988.
- [13] Njunda, A.L, Assob, J.C.N, Nsagha, D.S, Kamga, H.L, nde, P.F, and Yugah, V.C, Seroprevalence of toxoplasma gondii infection among pregnant women in Cameroon. *Journal of public health in Africa*, 2 (e24). 98-101. 2011.
- [14] Ogoina, D, Onyemelukwe, G.C, Musa, B.O, and Obiako, R.O, Seroprevalence of IgG and IgM antibodies to toxoplasma infection in healthy and HIV positive adults from Northern Nigeria. *Journal Infect dev tries*, 7 (5). 398-403. 2013.
- [15] Okwuzu, J.O, Odunukwe, N.N, Ezechi, O.C, Gbajabiamila, T.A, Musa, A.Z, Ezeobi, P.M, Oke, B.A, Somefun, T, Okoye, R, Onyeitu, C.C, Adetunji, M, and Otubanjo, A.O, Toxoplasma gondii infection in HIV/AIDS: prevalence and risk factors. *African Journal of clinical and Experimental microbiology*, 15 (2). 97-102. 2014.
- [16] Techalew, S, Mekashaw, T, Endale, T, Belete, T and Ashenafi, T, Seroprevalence of latent Toxoplasma gondii infection among HIV-infected and HIV-uninfected people in Addis Ababa, Ethiopia: A comparative cross sectional study. *BMC Research Notes*, 2: 13. 2009. [online] Available: <http://www.biomedcentral.com/1756-0500/2/213>. [accessed July. 31, 2014].
- [17] Verheij, R.A, Explaining urban-rural variations in health: A review of interactions between individual and environment. *Social Science and Medicine*, 42 (6). 923-35. 1996.
- [18] Wanachiwanawin, D, Sutthent, R, Chokephaibulkit, K, Mahakittikun, V, Ongrotchanakun, J and Monkong, N, Toxoplasma gondii antibodies in HIV and non-HIV infected Thai pregnant women. *Asia and Pacific Journal of allergy and immunology*, 19 (4). 291-293. 2001
- [19] Williams, C.C, and Windebank, J, helping people to help themselves: policy lessons from a study of deprived urban neighbourhoods in Southampton. *Journal of Social Science Policy*, 29 (3). 355-373. 2000.
- [20] Zemene, E, Yewhalaw, D, Abera, S, Belay, T, Samuel, A, and Zeynudin, A, Seroprevalence of toxoplasma gondii and associated risk factors among pregnant women in Jimma town, South Western Ethiopia. *BMC infectious diseases*, 12: 337. December 2012. [Online]. Available: <http://biomedcentral.com/1471-2334/12/337> [Accessed July. 25, 2014].
- [21] Zuffrey, J, Sugar, A, Rudaz, P, Bille, J, Glauser, M.P, and Chave, J.P, Prevalence of latent toxoplasma and serologic diagnosis of acute infection in HIV-positive patients. *European Journal of Clinical Microbiology and Infectious Diseases*, 12 (8). 591-5. 1993.