

Full Length Research Paper

The comparison of logistic regression models, on analyzing the predictors of health of adolescents, having multinomial response in Jimma Zone South-west Ethiopia

Girma Tefera*, Legesse Negash and Solomon Buke

Department of Statistics, College of Natural Science, Jimma University. Ethiopia.

Received 1 July, 2013; Accepted 20 May, 2014

Self-reported health status is the most commonly used measures of subjective and global measure of health because it is simple, economical and easy to administer. The objective of the study is to compare the performance of logistic regression models having multinomial response and identify the factors affecting health status of adolescents. Based on two stage sampling technique 2084 adolescents were interviewed to study the health status of teenagers in Jimma zone. In this article, we reviewed the most important logistic regression model and common approaches used to verify goodness-of-fit, using software R. We performed formal as well as graphical analyses to compare ordinal logistic regression models using data sets of health status. The results obtained from both baseline category logit model and ordinal logistic regression showed that sex of adolescents, source of drinking water and educational status significantly affect health status of teenagers. It was also found that a cumulative logit model containing these predictors provided the best description of the dataset among baseline category logit model, adjacent category logit model and continuation ratio model.

Key words: Adolescents' health status, multinomial logistic regression and ordinal logistic regression models, model comparisons using Akakie information criteria (AIC), goodness of fit.

INTRODUCTION

"Self-assessed", "self-reported" or "self-rated health" questions such as "How would you rate your current health status and would you say that it is very good, good, moderate/fair or poor/bad?" are among the most commonly used measure of subjective evaluation of

health status. Past studies have found this type of question to be a useful global measure of health (Zimmer et al., 2000).

The health status is usually classified as very good, good, moderate and poor/bad. When the researchers are

*Corresponding author. E-mail: girmastat@gmail.com. Tel: +251910185903.

interested in finding the determinants of self reported health status, usually two separate binary logistic regression models are required to develop by grouping the response variable into two categories. This task is tedious and cumbersome due to estimation and interpretation of more parameters.

In many epidemiological and medical studies, ordinal logistic regression model is frequently used when the response variable is ordinal in nature. The study has made an effort to identify the predictors of health status of adolescents using an ordinal logistic regression model and multinomial logit model and selecting the appropriate models among them.

The aim of the study is to compare the efficiency of multinomial logistic regression models and ordinal logistic regression models as well as identifying the significant predictors affecting self reported health status of adolescents.

MATERIAL AND METHODS

Baseline category logit (BCL) model

Even if the response is ordinal, we do not necessarily have to take the ordering into account. One category is arbitrarily chosen as the reference category. If it is the first category, then the logits for the other categories are defined by,

$$\text{logit}(\pi_j) = \log\left(\frac{\pi_j}{\pi_1}\right) = X_j^T \beta_j \text{ for } j = 2, 3, \dots, J$$

$$\hat{\pi}_j = \hat{\pi}_1 \exp(X_j^T \beta_j)_{\text{and}} \hat{\pi}_1 = \frac{1}{1 + \sum_{j=2}^J \exp(X_j^T \beta_j)}$$

Often, it is easier to interpret the effects of explanatory factors in terms of odds ratio than the parameters β . The odds ratio for exposure for response j ($j = 2, \dots, J$) relative to the reference category $j=1$ is,

$$\text{OR}_j = \frac{\pi_{jp} / \pi_{ja}}{\pi_{1p} / \pi_{1a}}$$

Where, π_{jp} and π_{ja} denote the probabilities of response category j ($j = 1, \dots, J$) according to whether exposure is present or absent, respectively.

Cumulative link models (CLM)

A cumulative link model is a model for an ordinal response variable, y_i that can fall in $j = 1, \dots, J$, categories. Then y_i follows a multinomial distribution with parameter π_{ij} , where π_{ij} denotes the probability that the i^{th} observation falls in the j^{th} response category. We define the cumulative probabilities as,

$$Y_{ij} = P(y_i \leq j | x) = \pi_{i1} + \dots + \pi_{ij}$$

$$Y_{ij} = \frac{\exp(\theta_j + X_i^T \beta)}{1 + \exp(\theta_j + X_i^T \beta)}$$

Where θ_j is the cut points or intercept for each logit and β is vector

of slopes for each logit. The CLM was originally proposed by Walker and Duncan (1967) and later called the proportional odds model by McCullagh (1980). The cumulative logits are also defined by Agresti (2002, 2007).

$$\text{logit}(Y_{ij}) = \text{logit}(P(Y_i \leq j)) = \log\left(\frac{P(Y_i \leq j)}{1 - P(Y_i \leq j)}\right)$$

$$= \log\left(\frac{\pi_{i1} + \pi_{i2} + \dots + \pi_{ij}}{\pi_{i(j+1)} + \dots + \pi_{iJ}}\right)$$

The odds ratio of the event $y \leq j$ at x_1 relative to the same event at x_2 is

$$\text{OR} = \frac{Y_j(x_1)/(1 - Y_j(x_1))}{Y_j(x_2)/(1 - Y_j(x_2))} = \frac{\exp(\theta_j + X_1^T \beta)}{\exp(\theta_j + X_2^T \beta)} = \exp((X_1^T - X_2^T)\beta)$$

This is independent of j . Thus the cumulative odds ratio is proportional to the distance between x_1 and x_2 which made McCullagh (1980) to call the proportional odds model (POM).

Adjacent Categories Model (ACL)

Another general model for ordered categorical data is the adjacent category model. As before, we let π_{ij} be the probability that individual i falls into category j of the dependent variable, and we assume that the categories are ordered in the sequence $j=1, \dots, J$. Now take any pair of categories that are adjacent, such as j and $j+1$. We can write a logit model for the contrast between these two categories as a function of explanatory variables:

$$\log\left(\frac{\pi_{ij}}{\pi_{i(j+1)}}\right) = \theta_j + X_i^T \beta_j \quad j = 1, 2, \dots, J - 1$$

Here, π_{ij} is i^{th} adolescence falls in j^{th} health rate category, $\pi_{i(j+1)}$ is the probability of i^{th} adolescent falls in $(j+1)^{\text{th}}$ health rate category.

Continuation Ratio Model (CRM)

Feinberg (1980) proposed an alternative method to the POM for the analysis of categorical data with ordered responses. The continuation ratio model can then be formulated as,

$$\log\left(\frac{P(Y = y_j | x)}{P(Y > y_j | x)}\right) = \theta_j + X_i^T \beta \quad j = 1, 2, \dots, J - 1$$

And could essentially be viewed as a ratio of the two conditional probabilities, $P(y = y_j | x)$ and $P(y > y_j | x)$. The odds ratio for continuation ratio for the k^{th} covariate x_k can be obtained directly from its model.

$$\text{OR} = \frac{P(Y = y_j | x_k^{(1)}) / P(Y > y_j | x_k^{(1)})}{P(Y = y_j | x_k^{(0)}) / P(Y > y_j | x_k^{(0)})} = \exp(\beta_k (x_k^{(1)} - x_k^{(0)}))$$

The proportional odds assumption

By proceeding with the model given by $(\text{logit}(y_{ij}) = \theta_j + x_i^T \beta)$ the assumption of the covariate effects are invariant to the cut points,

thus implying proportionality in the odds ratios. The proportional odds model can be considered as a series of $J-1$ binary logits where the β 's are constrained across the models such that: $\beta_1 = \beta_2 = \dots = \beta_{J-1} = \beta$.

Goodness of fit and deviance

The goodness of fit or calibration of a model measures how well the model describes the response variable. Assessing goodness of fit involves investigating how close values predicted by the model with that of observed values. The goodness-of-fit χ^2 process evaluates predictors that are eliminated from the full model, or predictors that are added to a smaller model. The question in comparing models is whether the log-likelihood decreases or increases significantly with the addition or elimination of predictor(s) in the model.

A more general measure called the deviance is defined for generalized linear models and contingency tables. The deviance is closely related to sums of squares for linear models (McCullagh and Nelder, 1989; Nelder and Wedderburn (1972). The deviance is defined as minus twice the difference between the log-likelihoods of a full (or saturated) model and a reduced model: $D = -2(\ell_{\text{reduced}} - \ell_{\text{full}})$

The full model has a parameter for each observation and describes the data perfectly while the reduced model provides a more concise description of the data with fewer parameters. A special reduced model is the null model which describes no other structure in the data than what is implied by the design. The corresponding deviance is known as the null deviance and analogous to the total sums of squares for linear models. The null deviance therefore also denoted the total deviance. The residual deviance is a concept similar to residual sums of squares and simply defined as: $D_{\text{resid}} = D_{\text{total}} - D_{\text{reduced}}$. A difference in deviance between two nested models is identical to the likelihood ratio statistic for the comparison of these models. Thus, the deviance difference, just like the likelihood ratio statistic, asymptotically follows a χ^2 -distribution with degrees of freedom equal to the difference in the number of parameters in the two models.

Model comparison with likelihood ratio tests

Model selection includes the choice of the type of model and variable selection within a model type. In this framework, the parameters estimating method with numerical integration has the advantage of being based on likelihood statistics. Thus, models can be ordered according to likelihood-based measures, such as Akaike's information criterion or Schwarz's Bayesian criterion (which judges a model by how close its fitted values tend to be the true expected values, as summarized by a certain expected distance between the two). In selecting a model, we should not think that we have found the "correct" one. Any model is a simplification of reality. However, a simple model that fits adequately has the advantages of model parsimony. If a model has relatively little bias, describing reality well, it provides good estimates of outcome probabilities and of odds ratios that describe effects of the predictors.

A general way to compare models is by means of the likelihood ratio statistic. If we consider two models, m_0 and m_1 , where m_0 is a sub-model of model m_1 , that is, m_0 is simpler than m_1 and m_0 is nested in m_1 . The likelihood ratio statistic for the comparison of m_0 and m_1 is $LR = -2(\ell_0 - \ell_1)$, where ℓ_0 is the log-likelihood of m_0 and ℓ_1 is the log-likelihood of m_1 . The likelihood ratio statistic measures the evidence in the data for the extra complexity in m_1 relative to m_0 . The likelihood ratio statistic asymptotically follows a χ^2 distribution with degrees of freedom equal to the difference in the number of parameter of m_0 and m_1 . The likelihood

ratio test is generally more accurate than Wald tests. Cumulative link models can be compared by means of likelihood ratio tests with the anova method. Here, AIC is used for model selection and

comparison. $AIC = -2\ell(\hat{\beta}) + 2P$, where $\ell(\hat{\beta})$ - the maximum log likelihood and p is the number of parameters. That is, a model having a smaller AIC value is the preferable model.

RESULT AND DISCUSSION

Results

The data comprise 2084 adolescents aged 13-17 years who were interviewed to study the health of adolescents in South west Ethiopia, Jimma zone. The adolescents' response was recorded on four ordinal scales (poor, moderate, good and very good). But counts for responses "poor" and "moderate" health rate are amalgamated into one category "poor/moderate" due to sparse cell counts (poor, 1.1% and moderate, 5.6%). From the total of 2084 adolescents, 81.2% had very good health status, 12.1% had good health status, and 6.7% had poor/moderate health status.

The significant variables in BCL model (using R package: MASS, R function: stepAIC) are used to determine a model with the minimum possible AIC (Akaike information criteria). Accordingly, sex, source of water and educational status are the selected variables to yield the minimum possible AIC of all the combinations. So, we fit the BCL model which consists of the variables that yield the minimum AIC, as the lowest AIC is the better fit (Table 1).

The maximum value of the log-likelihood function for the fitted model is -1241.5, giving the likelihood ratio chi-squared statistic $2(-1241.4 + 1260.5) = 38.1$. The statistic, which has 8 degrees of freedom (10 parameters in the fitted model minus 2 for the minimal model), is significant compared with the $X^2(8)$ distribution (p -value < 0.0001), showing also the overall significance of the model. That means the null hypothesis of all slope parameter is zero is rejected (at least one coefficient of the parameter is different from zero). The AIC value is $2503 = (-2*(-1241.5) + 2*10)$ for the above BCL model.

A difference in deviance between two nested models (Table 1) is identical to the likelihood ratio statistic for the comparison of these models (Holtbrugge and Schumacher, 1991). The deviance of the additive model which includes all covariates is 2467.8 and the deviance of the model which only includes the three predictors (that is, sex, source of water and education) is 2482.997. Therefore the likelihood ratio statistics which is $15.2 = (2482.997 - 2467.8)$, asymptotically follows a χ^2 -distribution with degrees of freedom the difference in the number of parameters of the two models, $22 - 10 = 12$ (that is, $X^2(12)$). Since the likelihood ratio statistics shows a p value of 0.23, it implies that we fail to reject the null hypothesis, $H_0: \beta_p = \beta_c = \beta_g = \beta_w = 0$ (the slope coefficients for place,

Table 1. Base line category logit model.

Predictors	Base line category logit model			
	log (π_2/ π_1)		log (π_3/ π_1)	
	Good vs. poor/moderate health status	Very good vs. poor/moderate health status		
	Estimate (SE)	OR (95% CI)	Estimate (SE)	OR (95% CI)
Intercepts (β_{0j})	-0.01(0.41)		1.14(0.33)	
Sex				
Female	Ref	1		1
male	-0.18(0.21)	0.84(0.55, 1.28)	0.41(0.18)	1.5 (1.06 , 2.13)
Source of water				
Unprotected	Ref	1		1
Tap or Protected	0.38(0.28)	1.46 (0.85 ,2.51)	0.55(0.22)	1.74 (1.12 ,2.70)
Educational status				
No schooling	Ref	1		1
Primary school	0.44(0.37)	1.56 (0.75 , 3.23)	0.81(0.30)	2.26(1.25 ,4.07)
Secondary or more	0.10(0.47)	1.11 (0.44 , 2.78)	0.33 (0.38)	1.39 (0.66 ,2.93)

SE=standard error of the estimate, OR=odds ratio and CI= confidence interval.

cooking place, workload and garbage disposal are zero). Therefore, the model which only includes the three predictors (that is; sex, source of water and education) is better than the model which includes all covariates. Likewise when we compare the BCL model of Tables 1 and Table 2, we are obtaining the likelihood ratio of 3.97 with 2 degrees of freedom having p value of 0.14. This also implies that the model which consists of the three predictors is better than the model having additional one predictor (Table 1). Besides, it has minimum AIC; it implies that the model including only the three predictors is the parsimonious model for the BCL model.

When we check the proportional assumption of CLM, after obtaining the possible combinations of covariates which reduce the AIC value, the score test for the proportional odds assumption is 4.3 which follows a χ^2 -distribution with degrees of freedom 4= (4 *(3-2)), that is $X^2(4)= 9.49$, having p values of 0.37. It implies that the proportional odds assumption is satisfied. And the likelihood ratio tests for ACL model and CRM for checking the proportional odds assumption are 6.27 and 4.4 having p values of 0.18 and 0.36 respectively. Therefore, there is no evidence against the proportional odds assumption. Hence, the proportional assumption holds for both models, so, we do not need to fit the non proportional odds model.

The second column of estimates in Table 2, for example, gives the log-odds of responding in category 1 ("poor/moderate") versus other categories ("good" and "very good"), the log-odds of responding in categories 1 and 2 ("poor/moderate" and "good") versus category 3

("very good"). The estimate of ACL model gives the log-odds of responding in category 1 ("poor/moderate") versus category 2 ("good") and category 2 ("good") versus category 3 ("very good"). The estimate of CRM gives the log odds of adolescents fall in one category of health status given the other better health status categories. Since the sign of the coefficients for a predictor is the same for all ordinal logistic regression models (Table 2), they have similar interpretations. For instance, the estimate of sex is -0.51, -0.31 and -0.49 for POM, ACL and CRM respectively. So the odd ratios of male adolescents for all models are less than one, implying that males have slightly better health than females.

The log-likelihood function for the CLM is -1243.6, giving the likelihood ratio chi-squared statistic $2*(-1243.6+1260.515) = 33.9$. The statistic, which has 4 degrees of freedom (6 parameters in the fitted model minus 2 for the minimal model), is significant compared with the $X^2(4)$ distribution (p-value < 0.0001), showing the overall significance of the model. That means at least one coefficient of the parameter is different from zero. For ACL model and CRM, the likelihood ratios are 31.76 and 33.68 with $X^2(4)$ respectively (p value <0.0001 for the two models), showing also the overall significance of the models. The likelihood ratio statistics of POM for the two nested models; that is, for fitted model on Tables 1 and 2 is 1.58, which asymptotically follows a χ^2 -distribution with degree of freedom 7 - 6 = 1 (that is, $X^2(1)$). Since the likelihood ratio statistics shows a p value of 0.21, it implies that we fail to reject the null hypothesis of $H_0: \beta_w$

Table 2. Ordinal logistic regression models for selected predictors.

Predictors	Ordinal logistic regression model								
	Proportional odds model			Adjacent category logit model			Continuation ratio model		
	Estimate (SE)	OR (95% CI)	P-value	Estimate (SE)	OR (95% CI)	P - value	Estimate (SE)	OR (95% CI)	P - value
Intercept₁ (θ_1)	-1.62 (0.24)			0.081 (0.18)			-1.68 (0.23)		
Intercept₂ (θ_2)	-0.43 (0.24)			-1.18 (0.169)			-0.93 (0.22)		
Sex									
Female	Ref	1			1			1	
Male	-0.51 (0.11)	0.6 (0.48,0.75)	< 0.001	-0.31 (0.08)	0.73 (0.63,0.86)	< 0.001	-0.49 (0.11)	0.61 (0.49,0.76)	<0.001
Source of water									
Unprotected	Ref	1			1			1	
Tap or Protected	-0.35 (0.15)	0.71 (0.52,0.95)	0.025	-0.25 (0.103)	0.78 (0.63,0.95)	0.013	-0.32 (0.15)	0.73 (0.55,0.97)	0.028
Educational status									
No schooling	Ref	1			1			1	
Primary school	-0.57 (0.21)	0.56 (0.37, 0.86)	0.008	-0.40 (0.14)	0.67 (0.51,0.88)	0.004	-0.54 (0.20)	0.58 (0.40,0.87)	0.007
Secondary (more)	-0.28 (0.27)	0.75 (0.45, 1.27)	0.290	-0.18 (0.18)	0.83 (0.59,1.18)	0.297	-0.27 (0.25)	0.77 (0.47,1.25)	0.290
Score test		4.2564			6.27			4.389	
Df.		4			4			4	
p-value		0.3724			0.18			0.356	
AIC		2499.171			2501.265			2499.35	

SE=standard error of the estimate, OR=odds ratio and CI= confidence interval.

=0, (the coefficient of workload). Therefore a model which excludes this variable is preferable than a model which includes it. Likewise, the likelihood ratio statistics of the two nested models for ACL model and CRM are 0.57 and 1.68 respectively, which follows a χ^2 -distribution with each degree of freedom $7 - 6 = 1$ (that is, $X^2(1)$). The likelihood ratio statistics shows a p value of 0.45 and 0.20; it implies just as the POM, we fail to reject the null hypothesis of $H_0: \beta_w = 0$ for ACL model and CRM. Generally, a model which is fitted using the three predictors (that is, sex, source of water and education) is better than a model which is fitted using the four univarily significant predictors or a model which includes all predictors for CLM, ACL model and CRM respectively. Having the maximum likelihood value for each model, it is possible to have their AIC value. Accordingly, the AIC value of POM is 2499.2, the AIC of ACL model is 2501.3 and the AIC of CRM is 2499.4.

Comparison of models

We used the likelihood ratio test to compare nested models, whereas AIC is used to compare the non-nested models. We compared all models using statistical criteria of log likelihood, goodness of fit and AIC. But choice of model should depend less on goodness of fit.

The ACL model corresponds to a BCL model. One can fit ACL by fitting the equivalent BCL model. But the construction of the ACL model recognizes the ordering of Y categories. To benefit from this model parsimony requires appropriate specification of the linear predictor. Since explanatory variable has similar effect for each logit, advantages accrue from having a single parameter instead of $2 = (3-1)$ parameters describing that effect. When used with this proportional odds form, ACL model fits well. Besides it has minimum AIC value as compared with BCL model.

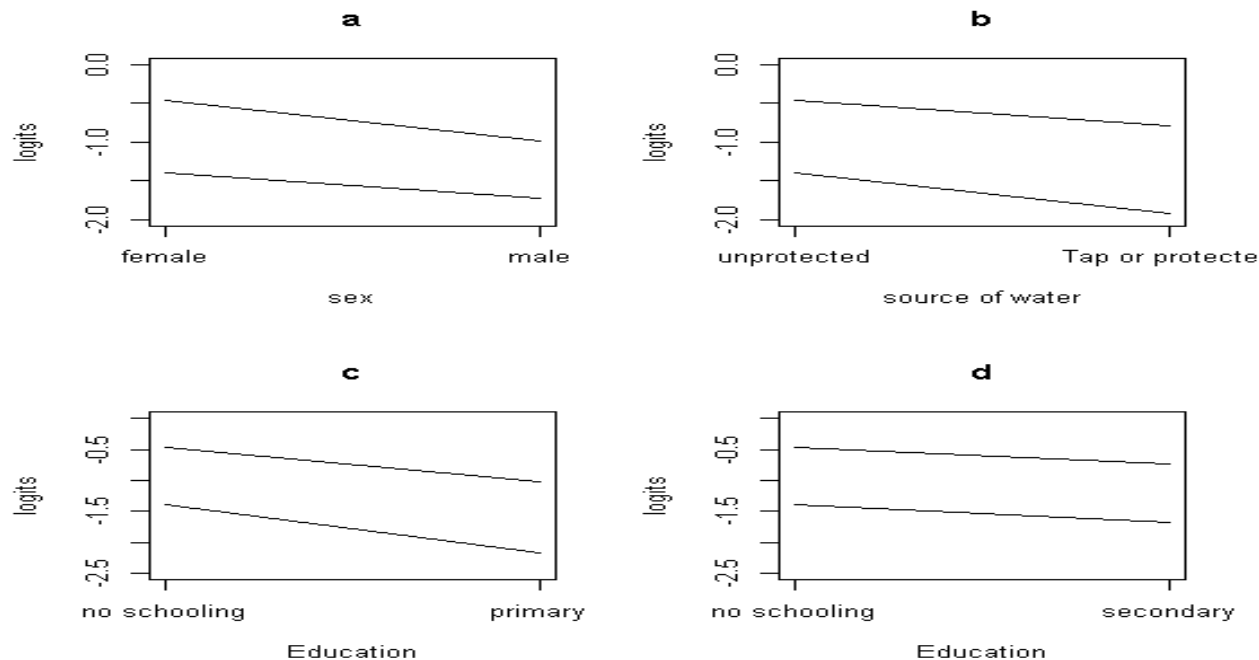


Figure 1. Proportional assumption of CLM.

Usually, the fit of both CLM and CRM is similar for many data sets. Here also the fit of the two models is almost similar. When we see the AIC values of the two models, the AIC of CLM is slightly smaller than that of CRM, has also slightly higher goodness of fit ($p=0.79$) than CRM ($p=0.78$) and proportionality is satisfied in better way than other models as its p value is the largest of all. Besides this, the CRM, is not invariant under an amalgamation of adjacent categories; for this reason, CRM is suitable in circumstances where the individual categories of the response are intrinsically of interest. So, CLM is better than CRM for this data set.

When we see the best model among the selected models, CLM fits well for this data set. It is also better than ACL model since it has minimum AIC value and goodness of fit for CLM has larger p value (0.79) than ACL model (0.65). Therefore, POM is the parsimonious model. Because, it satisfies the proportional assumption, has less number of parameters as compared with BCL model, shows model adequacy, has better goodness of fit and has smaller value of AIC as compared with the other models.

Generally, ordinal logistic regression model is better than nominal logistic regression model for this data set.

The final appropriate model is CLM that has two logits in which each logit is only different with their cut point values because of the fulfillment of the proportional odds assumption for this data set. The effects of the explanatory variables are the same across the three logit functions: $\text{logit}(\gamma_{ij}) = \theta_j - 0.51 \text{sex}_{\text{male}} - 0.35 \text{swater}_{\text{tap}} - 0.57 \text{educ}_{\text{primary}} - 0.28 \text{educ}_{\text{secondary}}$; Where, $\text{sex}_{\text{male}} = \text{male}$

adolescences, $\text{swater}_{\text{tap}} = \text{tap or protected source of water}$, $\text{educ}_{\text{primary}} = \text{primary education}$ $\text{educ}_{\text{secondary}} = \text{secondary education}$; $i = 1, 2, \dots, 2084$ and $j = 1, 2$ (Figure 1).

DISCUSSION

The POM and CRM are the most widely used in epidemiological and biomedical applications (Ananth and Kleinbaum, 1997) while other models for analysis of ordinal outcomes have received less attention. This is because both models may be interpreted in terms of odds ratios (familiar to epidemiologists), basic underlying assumption of each model—equality of β 's and statistical models may be plausible biologically. Armstrong and Sloan (1989) reported that usually both CLM and CRM are similar for many data sets. Here also the fit of the two models is almost similar in this study.

The POM can be viewed as a model nested with the unconstrained PPOM, and according to the deviance, the unconstrained partial proportion odds model is better than POM as it has a smallest p value ($p < 0.05$) and the proportionality assumption is violated (Ananth and Kleinbaum, 1997; Peterson and Harrell, 1990). But in this study POM is the selected model as the assumption is satisfied and had minimum AIC. Usually BCL models are better than ordinal logistic regression models when the proportional odds assumption is violated; in such cases BCL model can be treated as an alternative model for ordinal logistic regression model.

According to this study, CLM was found to be the better model than other models as it had minimum AIC, satisfied the proportional assumption and had better goodness of fit. Besides AIC, an intuitive choice between CLM and CRM can also be based on the goals of statistical analysis.

This finding is consistent with the results of other studies. For example; educational attainment was significantly associated with self-rated health, in the expected directions and females were slightly more likely than males to report fair or poor self-rated health (Veenstra, 2011).

Conclusion

Ordinal logistic regression models were better than nominal logistic regression model. Among ordinal logistic regression models the CLM or proportional odds model was an improved fit as compared to the rest models for any combination of variables in the data set. We also found that sex, source of drinking water and educational status of the adolescents had a significant effect on their health as they were the possible combinations to yield the minimum AIC in the CLM. Being literate and using of tap or protected water had a positive contribution for a better health status of teenagers but high workload which was univariately significant had a deteriorate impact on state of health and boys were less likely than females to report a deteriorate state of health.

Conflict of Interests

The authors have not declared any conflict of interests.

ACKNOWLEDGMENT

We want to thank Professor Tefera Belachew for his unreserved encouragements and help throughout the accomplishment of this work. We are grateful to Jimma Family Survey of Youth project which is supported by Brown University for permission to use the health status dataset for this study.

REFERENCES

- Agresti A (2002). *Categorical data analysis*, John Wiley & Sons, inc., 2nd edition. Hoboken, New Jersey.
- Agresti A (2007). *An introduction to categorical data analysis*, John Wiley & Sons, inc., 2nd edition. New Jersey.
- Armstrong BG, Sloan M (1989). Ordinal regression model for epidemiologic data. *Am. J. Epidemiol.* 129:191-204.
- Ananth CV, Kleinbaum DG (1997). "Regression Models for Ordinal Responses: Review of Methods and Applications." *Int. J. Epidemiol.* 26:1323-1333.
- Feinberg B (1980). *Analysis of cross classified data*, 2nd edition, Cambridge: Massachusetts institute of technology press.
- Holtbrugge W, Schumacher MA (1991). Comparison of regression models for the analysis of ordered categorical data, *Appl. Stat.* 40:249-259.
- McCullagh P, Nelder JH (1989). *Generalized linear models*. 2nd ed. London UK: Chapman and Hall.
- McCullagh P (1980). Regression models for ordinal data. *J. Royal Stat. Soc., Series B* (42):109-142.
- Nelder JA, Wedderburn RWM (1972). "Generalized Linear Models," *J. Royal Stat. Soc. Series A*, 135:370-384.
- Peterson BL, Harrell FE (1990). Partial Proportional odds model for ordinal response variable. *Appl. Stat.* 39:205-217.
- Walker SH, Duncan DB (1967). Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54:167-179.
- Zimmer Z, Natividad J, Lin H, Chayovan N (2000). A Cross-National Examination of the Determinants of Self-Assessed Health. *J. Health Soc. Behav.* 41:465-481.