J Bangladesh Agril Univ 18(2): 502–508, 2020 https://doi.org/10.5455/JBAU.82599

ISSN 1810-3030 (Print) 2408-8684 (Online)

Journal of Bangladesh Agricultural University

Journal home page: http://baures.bau.edu.bd/jbau

Analysis of overdispersed count data: A multilevel modeling approach

Fazle Elahi[1]✉, Soma Chowdury Biswas[2]

[1]Department of Agricultural and Applied Statistics, Bangladesh Agricultural University, Mymensingh 2202, Bangladesh
[2]Department of Statistics, University of Chittagong, Chittagong 4331, Bangladesh

ARTICLE INFO

Article history:
Received: 17 January 2020
Accepted: 26 March 2020
Published: 30 June 2020

Keywords:
Over-dispersed count data,
Multilevel model,
Antenatal care,
Maternal health services,
Bangladesh

Correspondence:
Fazle Elahi
✉: fazle.elahi1111@gmail.com

ABSTRACT

In this study, it is aimed to apply multilevel model with two levels in Poisson and Negative binomial regression models and to make comparison between these models to select a model which fits well the over-dispersed count data and finally, to identify the significant factors which influence the number of antenatal care visits of women during their pregnancy period. In this study, two mixed effect models (Poisson regression model with random effect and negative binomial regression model with random effect) are applied to a real data set to obtain the potential determinants of number of antenatal care (ANC) visits of women during pregnancy in Bangladesh, where data are extracted from Bangladesh Demographic and Health Survey (BDHS), 2014. The individual or within variation in each division is lower level (level-1) and between variation among the division is higher level (level-2). It is observed that between two mixed effect models-Negative Binomial regression model with random effect is selected as better model based on AIC, BIC and dispersion parameter for modeling the number of antenatal care visits of women in Bangladesh which is over-dispersed count data. Among the significant covariates, the place of residence, respondent's education, wealth index, respondent's husband's education, decision maker on respondent's health care and access to mass media are notable factors that are found highly associated with the number of antenatal care visits of women during their pregnancy period. Although both individual- and division-level characteristics have an influence on the inadequate and non-use of ANC, division-level factors have a stronger influence in the rural areas. The results suggest that for over dispersed count data, the negative binomial regression model with random effect is more suitable than Poisson. The results also suggest that much sensitization has to be done specifically in these rural areas to empower pregnant women and their husbands as to improve ANC attendance and utilization. Furthermore, health promotion programs need to increase consciousness about the importance of ANC visits during pregnancy in rural area to ensure the ANC visits among the rural women.

Introduction

Modeling count variable is a common task in micro econometrics, the social and political sciences. Poisson regression model is widely used in analysis of count data (Cameron et al., 1998). But the classical Poisson regression model for count data is often of limited use in these disciplines because empirical count data sets typically exhibit over-dispersion (Demetrio, J.H. 2007). In real life, the variance of count data is more than its mean, which violates the assumption of Poisson regression model. As a result, Poisson regression model is not appropriate in such case. To overcome this problem, it is appropriate to use negative binomial regression model (Richared Berk and John MacDonald, 2007). All of these models belong to the family of generalized linear models (GLMs, Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989). The negative binomial regression model can be used instead of Poisson regression model when it is considered that the data is over- dispersed (Gardner et al., 1995; Saffari and Adnan, 2011). To solve the over- dispersion problem the negative binomial (NB) regression model is used instead of the more conventional Poisson distribution (Richared Berk and John MacDonald, 2007).

Maternal mortality and infant mortality are the health indicators which show the greatest differential between the developed and developing countries. One of the main determinants of maternal and neonatal mortality is antenatal care visits during pregnancy. This study aims to investigate the selected factors affecting the number of antenatal care visits of women during their pregnancy period in Bangladesh using Bangladesh Demographic and Health Survey (BDHS) 2014 data. Since data on the antenatal care visits includes 0, 1, 2, 3,….. visits which is count data, this data leads to the count data model. Also BDHS 2014 data is multistage clustered data, so we applied multilevel model to estimate the true factor at different levels of hierarchy (Gelman, A. 2012).

Cite this article

Elahi, F. and Biswas, S.C. 2020. Analysis of overdispersed count data: A multilevel modeling approach. *Journal of Bangladesh Agricultural University*, 18(2): 502–508. https://doi.org/10.5455/JBAU.82599

In this study the units at lower level (level-1) are individuals who are nested within units at higher level (division: level-2). We have used two mixed effect models**:** Poisson regression model and negative binomial regression model with random effect which are also known as multilevel model with two levels, for the count data. To estimate these models, we have used maximum likelihood method, generalized estimating equation approach, Bayesian approach, etc (Simon J. Bond and Vernon T. Farewell. July, 2009). To compare the models, we used general linear mixed model approach through maximum linear estimation (MLE) approach (Peter F. Thall. 1998).

Materials and Methods

Data and variable

We used secondary data obtained from BDHS-2014.

In this study, the dependent variable is "Number of Antenatal Care Visits of Women in Bangladesh" which ranges 0 to 10 times of visits. On the other hand, in this study nine predictor variables are respondent's age, place of residence, division, source of drinking water, respondent's education, wealth index, respondents' husband's education, decision maker on respondent's health care and access to mass media. From these independent variables, we consider 'Division' as the level-2 variation or the random effect and the rest are fixed effect. To analyze the data we used two-level and multilevel regression models. Several types of tests, viz. goodness of fit, description of AIC (Makalic, D.F. November 22, 2008. Model selection Tutorial#1: Akaike's Information Criterion), BIC etc. are used to find the best model used in the study. In this study, R statistical software version R i386 3.3.2 (package lme4) are used.

The model and estimation procedures

Multilevel analysis is a suitable approach to take into account the social contexts as well as the individual respondents or subjects (Snijders, 2011). Normally these situations can be seen in the data collected by multi-stage stratified clustered sampling. The simplest and the most common multilevel model consider only two-level of analysis and this study deals only with this.

A multilevel model or a mixed model can be represented as,

$$Y_i = x_i^{\prime}\beta + Z_i + \varepsilon$$

Where, Y is known vector of observations, with mean $E(Y_i) = \theta_i = \log \mu_i = x_i'\beta$; x_i is the fixed effect vector of covariates; β is an unknown vector of regression coefficients of fixed effects; Z_i (i=1, 2,…,7) is the unknown random effects; ε is an unknown random errors, with mean E(ε) = 0 and variance, Var (ε) = R

Let, $\theta_i = x_i'\beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots\ldots + \beta_k x_k$

In this process, we consider a generalized linear model with link log,

i.e. we get, $\theta_i = x_i' \beta$

Poisson regression model with random effect

Poisson model with random effect is,

$$Y_i/\theta_i \sim Poisson\,(e^{\theta_i^*})$$

$$where,\quad \theta_i^* = \log \mu_i = x_i^{\prime}\beta + Z_i$$

$$\log \mu_i = \theta_i^* = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots\ldots\ldots\ldots\beta_p x_{pj} + Z_i$$

$$\mu_i = e^{\theta_i^*} = \exp(x_i'\beta + Z_i)$$

Here,

$\theta_i^* = x_i^{\prime}\beta + Z_i$; *where* Z_i (i = 1, 2,…,7) *is random effects and* $Z_i \sim N(0,\psi)$

Negative binomial regression model with random effect

In negative binomial regression with random effects the parameter μ_i is modeled

$$\log(\mu_i) = x_i^{\prime}\beta + Z_i \Rightarrow \mu_i = \exp(x_i^{\prime}\beta + Z_i)$$

Where β is the (p+1)×1 vector of unknown parameters associated with the known covariate vector $x_i' = (x_{i1}, x_{i2}, x_{i3}, \ldots x_{ip})$, where p is the number of covariates not including the intercept, and Z_i be the random effects which follows a multivariate normal distribution with mean zero and variance-covariance matrix ψ . i.e. $Z_i \sim MVN(0,\psi)$.

In our analysis the model is,

$$\log(\mu_i) = \beta_o + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \ldots.\ Z_i$$

Modeling the number of antenatal care visits of women in Bangladesh

Multilevel (with two-level) regression is given below:

glmer (factor (no. of antenatal care visit) ~ factor (respondent's age) + factor (place of residence) + factor (source of drinking water) + factor (respondent's education) + factor (wealth index) + factor (respondents husband's education) + factor (decision maker on respondent's health care) + factor (access to mass media) + (1| factor (division)).

Fitting different regression models with random effect for the number of antenatal care visits

Fitting the Poisson regression model with random effect

Using the number of antenatal care visits as dependent variable our proposed model is

$Log(\mu_i) = \beta_0 + \beta_{1j}X_{1j} + \beta_{2j}X_{2j} + \beta_{3j}X_{3j} + \beta_{4j}X_{4j} + \beta_{5j}X_{5j} + \beta_{6j}X_{6j} + \beta_{7j}X_{7j} + \beta_{8j}X_{8j} + Z_{ij}$

Where, the variables used in the model are defined as

X_1 = Respondent's age (3 categories): X_2 = Place of residence (2 categories); X_3 = Source of drinking water (3 categories); X_4 = Respondent's education (4 categories)
X_5 = Wealth index (3 categories); X_6 = Respondent's husband's education (4 categories); X_7 = Decision maker on respondent's health care (4 categories); X_8 = Access to mass media (2 categories).

All of these are fixed effects and Z_{ij} is the random effect or level-2 variation or cluster variation, where, Z_{ij} = Division with seven categories.

Results and Discussion

At first, we find whether there exists a significant association between the continuous variable, number of ANC visits and a categorical variable. The results of ANOVA are given in the Table 1.

Estimation of fixed parameters: (for Poisson regression model)

The Table 2 represents estimated parameters of the Poisson regression model for the number of antenatal care visits. The value of AIC of the above fitted model is 18658.6, BIC is 18780.0, residual deviance is 18620.6 on 4377 degrees of freedom following chi-square with 1 degree of freedom. The dispersion parameter is found to be 18620.6/4377=4.25419. The assumption of equal variance to the mean in Poisson distribution is violated since the dispersion parameter is greater than 1, an indication of over-dispersion in the data. This means that the parameters of the model have been over estimated and their standard errors have been under estimated which will not give a true reflection of the model.

Table 1. Mean number of ANC visits by the selected socioeconomic and demographic variables

Variables (n=Sample Size)	Mean ± SD	Percentage	95% CI	p-value
Respondent's age				0.000
Under 19 (915)	2.61 ± 2.313	20.8	(2.46, 2.76)	
19-29 (2629)	2.83 ± 2.484	59.8	(2.74, 2.93)	
Above 29 (852)	2.48± 2.349	19.4	(2.32, 2.64)	
Place of residence				0.000
Urban (1407)	3.56± 2.566	32	(3.43, 3.70)	
Rural (2989)	2.32 ± 2.252	68	(2.24, 2.40)	
Division				0.000
Barisal (522)	2.39±2.237	11.9	(2.19, 2.58)	
Chittagong (847)	2.57±2.429	19.3	(2.41, 2.74)	
Dhaka (779)	3.01 ±2.502	17.7	(2.83, 3.18)	
Khulna (506)	3.31±2.406	11.5	(3.10, 3.52)	
Rajshahi(533)	2.70 ±2.473	12.1	(2.49, 2.91)	
Rangpur (545)	3.21 ±2.434	12.4	(3.00, 3.41)	
Sylhet (664)	2.00±2.208	15.1	(1.83, 2.17)	
Source of drinking water				0.000
Other (806)	3.09 ± 2.572	18.3	(2.91, 3.27)	
Tap water (93)	3.57 ± 2.688	2.1	(3.02, 4.12)	
Tube-well water (3497)	2.61± 2.373	79.5	(2.53, 2.69)	
Respondent's education				0.000
Illiterate (587)	1.45 ± 1.815	13.4	(1.31, 1.60)	
Primary (517)	2.10 ± 2.269	11.8	(1.91, 2.30)	
Secondary (323)	3.62 ± 2.303	7.3	(3.37, 3.87)	
Above secondary (2969)	2.98 ± 2.463	67.5	(2.89, 3.07)	
Wealth index				0.000
Poor (1757)	1.83 ± 2.115	40.0	(1.73, 1.93)	
Middle (840)	2.46 ± 2.158	19.1	(2.31, 2.60)	
Rich (1799)	3.71 ± 2.465	40.9	(3.60, 3.82)	
Respondent's husband's education				0.000
Illiterate (1004)	1.79 ± 2.043	22.8	(1.67, 1.92)	
Primary (1331)	2.28± 2.314	30.3	(2.16, 2.41)	
Secondary (1393)	3.08 ± 2.335	31.7	(2.96, 3.20)	
Above secondary (668)	4.23 ± 2.501	15.2	(4.04, 4.42)	
Decision maker on respondent's health care				0.000
Respondent alone (520)	2.97 ± 2.511	11.8	(2.75, 3.18)	
Respondent and husband (2123)	2.83 ± 2.439	48.3	(2.73, 2.94)	
Husband alone (1417)	2.44 ± 2.380	32.2	(2.32, 2.57)	
Someone else (336)	2.80 ± 2.319	7.6	(2.55, 3.05)	
Access to mass media				0.000
No (1699)	1.78 ± 2.063	38.0	(1.68, 1.88)	
Yes (2727)	3.29 ± 2.454	62.0	(3.20, 3.39)	

Table 2. Estimated parameters of Poisson regression model

Independent variable	Categories	Estimated parameter	Standard Error	Odds ratio	Z value	Pr(>lzl)
Intercept		0.386068	0.074114		5.209	1.90e-07 ***
Respondent's Age	Under 19	…….	………	………	………	………..
	19-29	0.031875	0.023920	1.0324	1.333	0.182682
	Above 29	-0.007978	0.031329	0.9921	-0.255	0.798994
Place of residence	Urban	…..	…….	……	……..	…………
	Rural	-0.159515	0.020995	0.8526	-7.598	3.02e-14 ***
Source of Drinking water	Other	……...	……………	…………..	……………	……………
	Tap water	0.043635	0.059242	1.0446	0.737	0.461387
	Tube-well water	-0.020969	0.023454	0.9792	-0.894	0.371298
Respondent's education	Illiterate	……….	…………	…………	…………	………
	Primary	0.224886	0.047401	1.2522	4.744	2.09e-06 ***
	Secondary	0.416301	0.049798	1.5163	8.360	< 2e-16 ***
	Above secondary	0.367202	0.039642	1.4437	9.263	< 2e-16 ***
Wealth index	Poor	………	…………	……….	…………	………..
	Middle	0.107417	0.030329	1.1134	3.542	0.000398 ***
	Rich	0.328293	0.029206	1.3886	11.241	< 2e-16 ***
Respondent's husband's education	Illiterate	…………	…………….	………….	……………	………
	Primary	0.058598	0.031109	1.0603	1.884	0.059618.
	Secondary	0.151740	0.031906	1.1639	4.756	1.98e-06 ***
	Above secondary	0.323437	0.035691	1.3819	9.062	< 2e-16 ***
Decision maker about respondent's health	Respondent alone	……….	…………	…………...	…………	………
	Respondent & husband	-0.040414	0.028858	0.9604	-1.400	0.161378
	Husband alone	-0.110774	0.031111	0.8951	-3.561	0.000370 ***
	Someone else	-0.088591	0.041867	0.9152	-2.116	0.034343 *
Access to mass media	No	………..	………………..	……………	…………….	…………
	Yes	0.215419	0.026212	1.2404	8.218	< 2e-16 ***

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Estimation of random parameter (level-2 variation)

Groups Name	Variance	SD
Division (Intercept)	0.01593	0.1262

The estimate of the random intercept for level-2 (division) is 0.015, which means that the average variation of the number of antenatal care visits of women between divisions is 0.015.

Estimation of Fixed Parameters: (for negative binomial regression model)

The Table 3 represents estimated parameters of the Negative Binomial regression model for number of antenatal care visits. The value of AIC of the above fitted model is 17942.9, BIC is 18070.7 residual deviance 17902.9 on 4376 degrees of freedom following chi-square with 1 degree of freedom. The dispersion parameter is found to be 17902.9/4376= 4.09115. From the Table 3, it is observed that the parameter estimates have been increased and the standard errors have also increased. It is also observed that Negative binomial regression model reduced over-dispersion problem.

Estimation of Random parameter (level-2 variation)

Groups Name	Variance	SD
Division (Intercept)	0.02115	0.1454

The estimate of the random intercept for level-2 (division) is 0.02, which means that the average variation of the number of antenatal care visits of women between divisions is 0.02.

Model comparison and interpretation

The parametric comparison between two mixed effect models- Poisson and Negative Binomial regression models with random effect for AIC, BIC and dispersion parameter are given in Table 4. From the Table 4, it is clear that, the values of both AIC and BIC obtained from the real dataset are lowest in case of Negative Binomial regression model with random effect and the dispersion parameter for the Negative Binomial model with random effect is also smaller between two mixed effect models. Based on AIC, BIC and dispersion parameter it is found that between two mixed effect models, Negative binomial regression model with random effect is better model for modeling the number of antenatal care visits of women in Bangladesh which is over-dispersed count data.

This study found that negative Binomial regression model with random effect is better model for modeling the number of antenatal care visits of women in Bangladesh, The women in the selected model visit for antenatal care at least one time in their pregnancy period.

The intercept has been found to be 0.376866 in the Table 3 which is statistically significant at 0.1% level of significance. The intercept term 0.376866 represents that log of the average number of antenatal care visits of women during pregnancy who are under 19 and illiterate, comes from poor family in urban area of Bangladesh, whose husbands are uneducated, who drink other source of water, take the decision on their health care alone and have no access to mass media (Radio, Television and Newspaper) in the group of women who visit for antenatal care. In this case the expected number of antenatal care visits of above characterized women is exp(0.376866)=1.45771. That is, the above characterized women who visit for antenatal care take antenatal care 1.45771 times on average during their pregnancy period in Bangladesh. The estimated parameters corresponding to respondent's age are statistically insignificant. The estimated parameter of place of residence of rural category is -0.160040 which is statistically significant ($p < 0.01$). The women who live in rural area visit for antenatal care on an average 0.8521 times less than the women who live in urban area in the group of women who visit for antenatal care at least one time.

Table 3. Estimated parameters of negative binomial regression model

Independent variable	Categories	Estimated parameter	Standard error	Odds ratio	Z value	Pr(>|z|)
Intercept		0.376866	0.100835		3.737	0.000186 ***
Respondent's age	Under 19	……	…….	…….	……	……
	19-29	0.020538	0.037190	1.0208	0.552	0.580789
	Above 29	-0.024357	0.048742	0.9759	-0.500	0.617270
Place of residence	Urban	…….	…….	…….	……	……..
	Rural	-0.160040	0.033908	0.8521	-4.720	2.36e-06 ***
Source of Drinking water	Other	……	……	……	……	……
	Tap water	0.080005	0.099708	1.0833	0.802	0.422327
	Tube-well water	-0.002849	0. 037800	0.9972	-0.075	0.939918
Respondent's education	Illiterate	…….	…….	……	……	……
	Primary	0.224581	0.066486	1.2518	3.378	0.000730***
	Secondary	0.435050	0.074896	1.5450	5.809	6.29e-09 ***
	Above secondary	0.369339	0.054836	1.4468	6.735	1.64e-11 ***
Wealth index	Poor	……	……	……	…….	…….
	Middle	0.113738	0.044861	1.1205	2.535	0.011233 *
	Rich	0.338726	0.044822	1.4032	7.557	4.12e-14 ***
Respondent's husband's education	Illiterate	…….	…….	…….	…….	…….
	Primary	0.062354	0.044770	1.0643	1.393	0.163697
	Secondary	0.150528	0.047367	1.1624	3.178	0.001483 **
	Above secondary	0.324062	0.055663	1.3827	5.822	5.82e-09 ***
Decision maker about respondent's health	Respondent alone	…….	…….	……	…….	……..
	Respondent & Husband	-0.047363	0.046297	0.9537	-1.023	0.306299
	Husband alone	-0.131258	0.049247	0.8770	-2.665	0.007692 **
	Someone else	-0.102311	0.066397	0.9027	-1.541	0.123340
Access to mass media	No	…….	…….	…….	…….	…….
	Yes	0.224251	0.038442	1.2514	5.834	5.43e-09 ***

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 4. Parametric comparison between Poisson and Negative Binomial regression models with random effect

Source of model comparison	Poisson regression model with random effect	Negative Binomial regression model with random effect
Residual deviance	18620.60	17902.90
Degrees of freedom	4377	4376
Dispersion parameter	4.25	4.09
AIC	18658.60	17942.90
BIC	18780.00	18070.70

From the Table 3, it is found that primary, secondary and above secondary educational levels of respondents are highly significant. From the odds ratios it is clear that, the ratio of women who receive antenatal care and who never receive antenatal care is 1.2518, 1.5450 and 1.4468 times more among the primary, secondary and above secondary educated women respectively than the illiterate women. This means the percentage of women who never receive antenatal care is less among the educated women as compared to uneducated women. It also clear that the primary educated respondents visit for antenatal care on average about 25.2%, the secondary educated respondents visit for antenatal care on average about 54.5% and the above secondary educated respondents visit for antenatal care on average about 44.7% more than the illiterate respondents in the group of antenatal care visited respondents. The odds ratios for middle and rich levels of wealth index are 1.1205 and 1.4032 respectively, which are strongly associated with the number of visits for antenatal care of women during pregnancy period at 5% and 0.01% level of significance respectively. From this result it is clear that, the women of middle class family and rich family visit antenatal care on average respectively 1.1205 and 1.4032 times more as compared to the women of poor family.

The parameters corresponding to respondent's husband's education indicate that primary level of respondent's husband's education is statistically insignificant but secondary and above secondary both levels of respondent's husband's education are statistically significant at 1% level of significance and 0.01% level of significance respectively. The odds ratios secondary and above secondary levels of respondent's husband's education are 1.1624, and 1.3827 respectively. These indicates that the average number of antenatal care of visit of respondents whose husband is secondary and above secondary educated are respectively 1.1624, and 1.3827 times more than the respondents whose husband is illiterate in the group of respondents who at least one time take antenatal care.

The decision maker on respondent's health care is statistically significant. Only husband alone level of decision maker on respondent's health care is statistically significant at 1% level of significance. It is clear that the respondents whose health care decision is taken by husband alone they visit for antenatal care on average 0.8770 times less than the respondents who take the decision on their health care alone among the antenatal care taken respondents. Access to mass media is statistically significant at 0.01% level of significance. This indicates access to mass media has impact on the number of antenatal care visit. From the analysis it is found that, the odds ratio of access to mass media is 1.2514 which means the average number of antenatal care visits of women who have access to mass media about 25.1% more than women who have no access to mass media in the group of antenatal care visited women.

The negative binomial regression model with random effect is presented in the following fitted equation:
Log (mean number of visits) $= 0.376866 + 0.020538 X_{12} - 0.024357 X_{13} - 0.160040 X_{22} + 0.080005 X_{32} - 0.002849 X_{33} + 0.224581 X_{42} + 0.435050 X_{43} + 0.369339 X_{44} + 0.113738 X_{52} + 0.338726 X_{53} + 0.062354 X_{62} + 0.150528 X_{63} + 0.324062 X_{64} - 0.047363 X_{72} - 0.131258 X_{73} - 0.102311 X_{74} + 0.224251 X_{82} + Z_{ij}$

Where, X_{1j}'s represent the levels of respondent's age, X_{2j}'s represent the level of place of residence,X_{3j}'s represent the level of source of drinking water, X_{4j}'s represent the level of respondent's education, X_{5j}'s represent the level of wealth index categories of the respondent's, X_{6j}'s represent the level of respondent's husband's education X_{7j}'s represent the level of decision maker on respondent's health care and X_{8j}'s represent the level of access to mass media of the respondent's and Z_{ij} is the level-2 variation or cluster variation of the model, where j for individual (Level-1), and i for division (Level-2).

Discussion

The variance of the random intercept term, which shows the extent to which outcomes between divisions differ, after controlling for the covariates. The estimate of the random intercept for level-2 (division) is 0.02, which means that the average variation of the number of antenatal care visits of women between division is 0.02. Although both individual- and division-level characteristics have an influence on the inadequate and non-use of ANC, division-level factors have a stronger influence in the rural areas.

In this study it was founded that the estimated parameter of place of residence, the educational levels of respondent all are highly significant. This result is consistent with the findings of Islam and Odland (2011). In this study it is founded that the wealth index which are strongly associated with ANC, and this finding is supported by that of a study by Shahjahan *et al*. (2012) which found that wealth index had a significant association with the use of ANC services.

We found that the parameters corresponding to respondent's husband's education is statistically significant with ANC visits. This is also consistent with the research paper, Contextual Influences on the Use of Antenatal Care in Nepal, by- Matthews, Stephen A., and Bina Gubhaju.(2004) and Utilization of focused antenatal care in Zambia: Examining individual- and community-level factors using a multilevel analysis, by Chiliba and Koch (2013).

It was found that the decision maker on respondent's health care is statistically significant. This finding is supported by Matthews, Stephen A., and Bina Gubhaju. (2004) which found that there is significant association of decision maker about respondent's health care with the use of ANC services. The results of this study showed that the access to mass media is statistically significant with the number of antenatal care (ANC) visits. This finding is supported by and also consistent with the results of research by-Priyanka Dixit *et al*. (2013), Islam and Odland (2011) and Shahjahan *et al.* (2012). We found that several studies from Bangladesh and other countries have demonstrated that woman's education, place of residence, wealth index, respondent's husband's education, decision maker about respondent's health care and access to mass media are the most important determinant of ANC utilization, and in this study these are also the significant determinants of antenatal care (ANC) visits. Although all the results of the relevant literature are similar but most of them are analyzed using single level model or fixed effect model. On the contrary, in this study multilevel model with two level is used which gives more robust estimator.

Conclusion

In this study it is found that the multilevel effects (division level) are significant and have to take into consideration in mixed effect model which leads multilevel analysis. The study provides evidence that, while both individual and division-level factors are instrumental in determining the attendance and utilization of ANC. The results suggest that for over dispersed count data, the negative binomial regression model with random effect is more suitable than Poisson. Based on findings of this study, we can say that the women who have secondary educational qualification, come from rich family, live in urban area of Bangladesh, whose husband's educational qualification is above secondary, take the decision on their health care alone and have access to mass media (Radio, Television and Newspaper) visit more times for antenatal care among the women who visit for antenatal care, whereas in the class of women who have no educational qualification, come from poor family, live in rural area of Bangladesh whose husbands are illiterate and have no access to mass media. So we may conclude that this study can help policymakers and program managers have to track the progress of mothers' health and refocus efforts to meet the goal of reducing maternal and child mortality and morbidity to a great extent.

References

Bangladesh Demographic and Health Survey (BDHS). 2014. NIPORT, Dhaka, Bangladesh; Mitra and Associates, Dhaka/ Bangladesh.

Berk, R. and MacDonald, J. 2007. Overdispersion and Poisson Regression-Ensemble method for Data Analysis in the Behavioral, Social and Economic Science. *National Science Foundation.*

Cameron, A. C. and Trivedi, P.K. 1998. Regression Analysis of Count Data. *Cambridge: Cambridge University Press.*

Chiliba, C., Koch, S. 2013. Utilization of focused antenatal care in Zambia: Examining individual- and community-level factors using a multilevel analysis. Published by Oxford University Press in association with The London School of Hygiene and Tropical Medicine. *Health Policy and Planning,* 30:78–87

Demetrio, J.H. 2007. *Overdispersion: Model & Estimation.*

Gardner, W., Mulvey, E.P. and Shaw, E.C. 1995. Regression analysis for counts and rates: Poisson, Over-dispersed Poisson and Negative Binomial models. *Psychological Bulletin,* 118: 392-404.

Gelman, A. 2012. Multilevel (hierarchical) modeling: what it can and cannot do. *Technometrices.*

Islam, M.R. and Odland, J.O. 2011. Determinants of antenatal and postnatal care visits among Indigenous people in Bangladesh: a study of the Mru Community. *Rural and Remote Health,* 11: 1672.

Makalic, D.F. November 22, 2008. Model selection Tutorial#1:Akaike's Information Criterion. In Model selection with AIC. Melbourne.

Matthews, Stephen A. and Bina Gubhaju. 2004. *Contextual influences on the use of antenatal care in Nepal.* DHS Geographic Studies 2. Calverton, Maryland USA: ORC Macro.

McCullah, P. 1983 and Nelder, J.A. 1989. *Generalized Linear Models.* 2nd edition, Chapman and Hall, London, UK.

Nelder, J. and Wedderburn, R. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society,* Vol.-135(3): 370-384.

Peter F. Thall. 1998. Mixed Poisson Likelihood Regression Models for Longitudinal Interval Count Data. *International Biometric Society*, 44:197-209.

Priyanka Dixit, Laxmi Kant Dwivedi and Faujdar Ram. 2013. Estimating the impact of antenatal care visits on institutional delivery in India: A propensity score matching analysis. *Healt,* 5: 862-878. http://dx.doi.org/10.4236/health.2013.55114

Saffari, S.E., Adnan, R. and Greene, W. 2011. Handling of Over-Dispersion of Count Data via Truncation using Poisson Regression Model. *Journal of Computer Science & Computational Mathematics,* 1(1).

Shahjahan, M., Chowdhury, H.A., Akhter, J., Afroz, A., Rahman, M. M. and Hafez, M.A. 2012. Factors associated with use of antenatal care services in a rural area of Bangladesh. *South East Asia Journal of Public Health,* 2(2):61-66.

Simon J. Bond and Vernon T. Farewell. July, 2009. Likelihood Estimation for a Longitudinal Negative Binomial Regression Model with Missing Outcomes. *Journal of the Royal Statistical Society. Series C, (Applied Statistics)*, 58(3): 369-382

Snijders, T. A. 2011. *Multilevel analysis.* Springer.