Analyzing Overdispersed Antenatal Care Count Data in Bangladesh: Mixed Poisson Regression with Individual–Level Random Effects

Zakir Hossain
University of Dhaka, Bangladesh

Maria
University of Dhaka, Bangladesh

Abstract

Poisson regression (PR) is commonly used as the base model for analyzing count data with the restrictive equidispersion property. However, overdispersed nature of count data is very common in health sciences. In such cases, PR produces misleading inferences and hence give incorrect interpretations of the results. Mixed Poisson regression with individual–level random effects (MPR_ILRE) is a further improvement for analyzing such data. We compare MPR_ILRE with PR, quasi-Poisson regression (Q_PR) and negative binomial regression (NBR) for modelling overdispersed antenatal care (ANC) count data extracted from the latest Bangladesh Demographic and Health Survey (BDHS) 2014. MPR_ILRE is found to be the best choice because of its minimum Akaike information criterion (AIC) value and the overdispersion exists in data has also been modelled very well. Study findings reveal that on average, women attended less than three ANC visits and only 6.5% women received the World Health Organization (WHO) recommended eight or more ANC visits for the safe pregnancy and child birth. Administrative division, place of residence, birth order, exposure of media, education, wealth index and body mass index (BMI) have significant impact on adequate ANC attendance of women to reducing pregnancy complications, maternal and child deaths in Bangladesh.

Keywords: overdispersion, individual–level random effects, Akaike information criterion, World Health Organization.

1. Introduction

In many practical fields, such as public health, epidemiology, insurance, demography, psychology and actuarial studies, count data are often a very common phenomenon. Initially for analyzing these data, PR is used as the base or standard model with the equality assumption of mean and variance of count responses. However, in many real count data analyses, this assumption is often violated because of greater variability i.e., higher variance than mean. This is well-known as overdispersion. Ignorance of the consequences of overdispersion in modelling count data gives misleading inferences and interpretations about regression parameters. Thus, in case of analyzing count data, one should take much care to handle the overdispersion, otherwise estimates obtained from model fitting may not be precise and consistent. Q_PR and NBR models may be alternatives to deal with overdispersed nature of count data instead of

PR model (Hilbe 2011; Ver Hoef and Boveng 2007). Moreover, individual-level random effect (ILRE) models the extra-Poisson variation present in the data effectively via generalized linear mixed models (GLMMs) framework using a single-level of the random effect for each data value (Harrison 2014).

Maternal health care during pregnancy is an important focused issue for the development of a country. In many developing countries, like Bangladesh, millions of women experience a lot of life-risk difficulties during their pregnancy. Pregnancy and childbirth complications are the major causes of deaths and disabilities among women in Bangladesh (Chowdhury, Islam, Chakraborty, and Akhter 2007; Latif, Hossain, and Islam 2008; Mohammad, Zahura, and Rahman 2017; Islam and Sultana 2019). While such complications are very common among pregnant women in Bangladesh, very limited attention or effort has been given by the concerned authorities and researchers in this area to reduce these life-threatening complications by confirming satisfactory ANC attendance during the whole course of pregnancy.

About 810 maternal deaths were reported worldwide every day in 2017 due to major complications experienced by women during their pregnancy and childbirth. Unfortunately, 94% of all these maternal deaths occurred in developing countries (WHO 2019). Between 2000 and 2017, the maternal mortality ratio (MMR), defined by number of maternal deaths per 100,000 live births, declined by about 38% globally. In 2017, the estimated MMR was 462 maternal deaths in developing countries whereas this value was found only 11 for developed countries (WHO 2019). However, in developing countries, like Bangladesh the situation has not enriched throughout this time period. In current years, the MMR is observed to be steady in Bangladesh, because the estimated value of MMR in 2016 was 196 maternal deaths, nearly identical as the estimated MMR found in 2010 (NIPORT, ICDDRB, MEASURE Evaluation 2017). By 2030, one of the main targets of the sustainable development goals (SDGs)-3 is to drop the MMR to below 70 maternal deaths (WHO 2018) and this is undoubtedly a huge challenging problem.

ANC assists mother to be prepared for the risk-free delivery and childbirth by reducing pregnancy-and-delivery associated major complications. The utilization of ANC services, especially adequate ANC visits play an important role to reducing such complications and hence declining maternal and child deaths (Pandit 1992; Titaley, Hunter, Heywood, and Dibley 2010). WHO (2007) suggested that women should receive at least four ANC visits throughout the pregnancy to confirm their safe motherhood under usual conditions. However, this requirement was not sufficient to reducing pregnancy complications effectively and hence to give safe childbirth. As a result WHO (2016) updated and further recommended at least eight ANC visits for women during the whole course of pregnancy: one visit within first three months (first trimester) of pregnancy, two visits between three to six months (second trimester), and five visits in the third trimester (WHO 2016). However, the current scenario in Bangladesh reflects that women do not ensure this minimum target of ANC visits for their safe motherhood (NIPORT, Mitra & Associates, ICF International 2016).

The utilization of ANC attendance among women throughout the pregnancy is associated with different socio-economic and demographic variables. The frequency of ANC visits of women was highly affected by their age at first birth (Nisar and White 2003). The rate of ANC visits of pregnant women was greater among higher-economic groups than others (Jayaraman, Chandrasekhar, and Gebreselassie 2008). Women took significantly higher ANC visits during their first birth of pregnancy than the second or above births (Navaneetham and Dharmalingam 2002). Women living in urban areas were more likely to attend ANC visits than women from rural areas in Bangladesh (Rahman 2009). Women education was also found to be the significant factor for ensuring adequate ANC visits, antenatal morbidity, and also to reducing child and maternal deaths (Becker, Peters, Gray, Gultiano, and Black 1993; Mohammad *et al.* 2017; Chakraborty, Islam, Chowdhury, Bari, and Akhter 2003; Chowdhury *et al.* 2007). A study on ANC attendance among Bangladeshi women, who started to visit health centre during their pregnancy in order to receive antenatal care, reported that the urban and educated women who belonging to rich family and having access to mass–media

were more likely to attend adequate ANC visits compared with their counterparts (Hossain, Akter, Sultana, and Kabir 2020).

It is evident from the latest BDHS 2014 that Bangladeshi women do not complete the required minimum standard number of ANC visits, recommended by WHO (2007, 2016), for their safe pregnancy and childbirth (NIPORT, Mitra & Associates, ICF International 2016). A number of studies have been conducted to identify the associated risk factors for the utilization of adequate ANC attendance in Bangladesh (Sultana and Bari 2017; Islam and Masud 2018a,b; Rahman and Hossain 2019; Bhowmik, Das, and Islam 2020). The consequence of overdispersion exists in ANC count data of women in Bangladesh, extracted from BDHS 2014, was ignored in most of the previous studies while modelling such data. However, this has been considered in the present study to avoid misleading inferences and interpretation of results. Therefore, overdispersion has been detected and tested to investigate whether it is statistically significant or not in the data set used in this paper. We then aim to analyze the data adopting suitable statistical modelling in order to explore the possible significant factors for the WHO recommended satisfactory number of ANC attendance among women to receiving antenatal health care during pregnancy in Bangladesh.

2. Materials and methods

2.1. Data and sampling design

We used the latest survey data on ANC among women during pregnancy in Bangladesh extracted from BDHS 2014 (NIPORT, Mitra & Associates, ICF International 2016). In this survey, a two-stage stratified random sampling design was used in order to collect data. In the first stage, 600 enumeration areas (EAs) were considered with probability proportional to the size of EAs where 207 and 393 EAs are from urban and rural areas, respectively. A sampling frame of households in the selected EAs was then prepared. On average, 30 households were taken from each EA in the second stage of sampling. Finally, 18000 households were considered for the interview in this survey. The detailed description of data and survey can be found at the website: https://dhsprogram.com/data/available-datasets.cfm.

2.2. Variables

In this study, we aim to find out the potential determinants for adequate ANC attendance, recommended by WHO, among women during their pregnancy in Bangladesh. The number of ANC visits of Bangladeshi women to receiving health care during their pregnancy is considered as the count response variable. Administrative division (region), place of residence, birth order, mother's age at birth, exposure of media, education, wealth index, membership of non-government organization (NGO), women empowerment, education gap between husband and wife, religion, mother having son and BMI are considered as explanatory variables.

Information on all of the selected variables were not found straightforward from the survey data. For example, the variables: exposure of media, NGO membership and women empowerment are prepared by combining the related covariates. The women who read magazines or newspapers or watch television or listen to radio at least once per week are considered as the exposed group to mass media. If women who were involved with any one of the NGO's: Grameen Bank, Proshika, Bangladesh Rural Development Board (BRDB), Bangladesh Rural Advancement Committee (BRAC) and Association of Social Advancement(ASA) then they were considered to have the NGO membership otherwise not. If women were associated with any one of the decisions such as: buying major household goods, own and children health care, and visiting relatives or family then women were treated to be empowered. A data set of 4390 women information is finally obtained after deleting all the missing cases for this study.

2.3. Overdispersion in count data analysis

Overdispersion is a common phenomenon for modelling count response data which arises due to the excess variability i.e., variance of the responses is greater than the mean in a PR model. This is the case of ANC attendance count data considered in this study. One should take much care about the consequence of overdispersion in analyzing count data as the ignorance of this gives incorrect interpretations of the model parameters.

Detection

One may detect the overdispersion at the first stage whether it is present or not in a real data set by using Pearson chi-square (χ^2) statistic (Hilbe 2011), defined by

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)}, \tag{1}$$

where V is the variance function and $\hat{\mu}_i$ is the expected counts. If the *dispersion* value, Pearson residual χ^2 statistic divided by the corresponding degrees of freedom, is greater than 1 then the model is considered as overdispersed, while the value of *dispersion* is 1 for equidispersed; and less than 1 for underdispersed.

Tests

The amount of overdispersion occurred in the data, whether it is statistically significant or not, can be assessed by both the score and Lagrange multiplier tests. Three different versions of the score test are given by Dean and Lawless (1989):

$$Z_i = \frac{(y_i - \hat{\mu}_i)^2 - y_i}{\hat{\mu}_i \sqrt{2}}, \tag{2}$$

Cameron and Trivedi (1990):

$$Z_i = \frac{(y_i - \hat{\mu}_i)^2 - y_i}{\hat{\mu}_i} \tag{3}$$

and Winkelmann (2008):

$$Z_i = \frac{(y_i - \hat{\mu}_i)^2 - y_i}{2\hat{\mu}_i}. \tag{4}$$

Moreover, the statistic for Lagrange multiplier test (Hilbe 2011) of overdispersion can also be written as

$$L = \frac{\sum_{i=1}^{n}(\hat{\mu}_i^2 - n\bar{y}_i)^2}{2\sum_{i=1}^{n}\hat{\mu}_i^2}. \tag{5}$$

2.4. Models

We consider three different generalized liner models (GLMs): PR, Q_PR and NBR (McCullagh and Nelder 1989; Dobson and Barnett 2018) and an extended GLMM: MPR_ILRE (Harrison 2014) in order to select the best choice for analyzing overdispersed ANC count data of women in Bangladesh. PR is the first step for modelling count data with the equality assumption of mean and variance of the responses. Q_PR and negative binomial regression (NBR) are alternatives as both models incorporate the overdispersion for analyzing overdispersed count data. Let Y be a count response variable. It follows that the expectation of Y is $E(Y) = \mu > 0$. For the Q_PR model (Ver Hoef and Boveng 2007)

$$var(Y) = V(\mu) = \theta\mu, \tag{6}$$

where $var(Y)$ is the variance of Y, V is the variance function as before and $\theta > 1$ is the overdispersion parameter However, in case of NBR model, the variance of Y can be expressed as

$$var(Y) = V(\mu) = \mu + \kappa\mu^2 = \mu(1 + \kappa\mu), \tag{7}$$

where $\kappa > 0$. The overdispersion i.e., the excess amount of μ in (7) is $(1 + \kappa\mu)$, multiplicative factor, which depends on μ. It is to be noted that the variance function V is a linear function of μ for the Q_PR whereas this is a quadratic function for the NBR model.

Let the mean $\mu_i > 0$; $i = 1, \ldots, n$ for the ith woman vary as a function of covariates for that woman and it can be modelled in the context of GLMs framework (McCullagh and Nelder 1989; Dobson and Barnett 2018) as

$$g(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}, \tag{8}$$

where g and η_i are the log-link function and i^{th} linear predictor, respectively. Equation (8) can be rewritten in matrix notation as

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta}, \tag{9}$$

where $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{X}$ is the $n \times p$ design matrix with $p = k+1$, $\boldsymbol{x}_i = (1, x_{i1}, x_{i2}, ..., x_{ik})^\top$ is the ith row vector of $\boldsymbol{X}$, $\boldsymbol{\eta}$ is the vector of linear predictors and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, ..., \beta_k)^\top$ is the $p \times 1$ vector of regression coefficients. For the further improvement of modelling count data, one may consider MPR_ILRE in the context of GLMMs (Stroup 2012) by introducing random effects in the linear predictor of (9) as

$$g(\boldsymbol{\mu}) = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \tag{10}$$

where $\boldsymbol{\mu} = E(\boldsymbol{Y}|\boldsymbol{u})$ is now the conditional mean vector of responses given random effects, $\boldsymbol{Y}$ is a $n \times 1$ vector of responses and $\boldsymbol{u}$ is a $n \times 1$ vector of individual–level random effects. The components of $\boldsymbol{u}$ are assumed to be uncorrelated and normally distributed for the computational simplicity. We implement maximum likelihood estimation algorithm using the Laplace approximation for model parameters in the MPR_ILRE via the R–package: glmmTMB (Magnusson, Skaug, Nielsen, Berg, Kristensen, Maechler, Bentham, Koen, Sadat, Bolker, and Brooks 2020). We use AIC as the model selection criterion and can be written as (Akaike 1998)

$$AIC = -2l(\hat{\boldsymbol{\theta}}; \boldsymbol{y}) + 2p, \tag{11}$$

where p is the number of regression parameters to be estimated, $l(\hat{\boldsymbol{\theta}}; \boldsymbol{y})$ is the log-likelihood, $\hat{\boldsymbol{\theta}}$ is the vector of estimated model parameters and $\boldsymbol{y}$ is the data vector. A model with the smallest AIC value is considered as the best choice for analyzing the overdispersed ANC count data of women in Bangladesh. It is useful to consider incidence rate ratio (IRR) instead of regression of coefficients for interpretations and better understanding of results in order to examine the impact of explanatory variables on the count response variable. For the specific covariate x_j; $j = 1, \ldots, k$ the IRR is estimated as $\text{IRR}_j = e^{\hat{\beta}_j}$ where $\hat{\beta}_j$ is the jth estimated regression coefficient.

3. Results and discussion

Table 1 shows the frequency and percentage distribution of ANC visits of total 4390 women during their last pregnancy three years prior to the survey date. It is observed that 21.3% women did not visit any health centre to receive ANC during their whole period of pregnancy. According to WHO (2007) guideline, it is the minimum standard requirement for the pregnant women to have at least four ANC visits for the safe motherhood, whereas this figure was 32.2% for Bangladeshi women during their pregnancy. It can also be seen that a significant portion of participants i.e., 93.5% women visited less than eight times to a medically trained ANC health centre and hence, it follows that only 6.5% women attended the WHO (2016) further recommended eight or higher ANC visits in order to give the safe pregnancy and childbirth. Moreover, the women took on average only less than three (mean=2.86) ANC visits. The variance (14.79) of ANC count responses was also found to be substantially higher than its mean.

Table 1:   Frequency and percentage distribution of number of ANC visits with descriptive statistics among women aged 15-49 years in Bangladesh.

Distribution	Number of ANC visits of women									Total
	0	1	2	3	4	5	6	7	8+	
Frequency	935	711	727	602	482	296	183	167	287	4390
Percentage (%)	21.3	16.2	16.6	13.7	11.0	6.7	4.2	3.8	6.5	100

Note: mean=2.86, variance=14.79 (greater variability)

The percentage distributions of women among different categories of socio–economic and demographic variables are presented in Table 2. Out of seven administrative divisions in Bangladesh, 51.9% of women were from three major divisional cities, namely Chittagong (19.2%), Dhaka (17.7%) and Sylhet (15.0%). The percentage of women was very close for other regions as Barisal (11.9%), Khulna (11.7%), Rajshahi (12.2%) and Rangpur (12.3%). Most of the participants were muslim women (91.9%), living in rural areas (67.9%) and empowered to give opinion for important decisions in their family (80.3%). The percentages of women having media exposure (62.2%), higher education gap (82.3%), son (67.8%) and not having any NGO membership (67.9%) were substantially higher than their counterparts. Most of the women were below or equal to aged 35 years (95.9%) and 31.3% of them were aged below 20 years; had the third birth or below (86%) where 40.5% had the first birth during the survey period. Out of total respondents, 13.3% women had no educational attainment and 86.7 % had at least primary level of education. About 40% of women were from the poor family and 41.4% mothers had either under–weight (24.5%), over–weight (14.1%) or obesity (2.8%).

The analysis of variance (ANOVA) test was performed to examine whether there exists any significant variation among average ANC visits received by women of different associated groups of the covariates. The summary results are given in Table 2. From p–values in Table 2, it is observed that there was a highly significant association between average number of ANC visits and place of residence ($p<0.001$), birth order ($p<0.001$), exposure of media ($p<0.001$), education ($p<0.001$), wealth index ($p<0.001$), women empowerment ($p=0.002$) and BMI ($p<0.001$) of pregnant women in Bangladesh. It can also be seen that administrative division ($p=0.012$) and mothers having son ($p=0.044$) were significantly associated with mean ANC visits of women at 2% and 5% level, respectively. These nine significant covariates have been considered later in adjusted multivariate analysis for investigating potential determinants affecting ANC attendance among women during pregnancy in Bangladesh. From Table 2, it is seen that the highest mean number of ANC visits was 3.64 for women of Khulna region and this value belongs to 95% confidence interval (CI):3.21–4.07. On the other hands, the lowest average number of ANC visits (mean=2.03, 95% CI: 1.86–2.21) was found among women in Sylhet region. It can also be seen that women living in urban areas of Bangladesh experienced higher average number of ANC visits (mean=3.82, 95% CI:3.58–4.05) than rural areas (mean=2.42, 95% CI:2.29–2.54). The average ANC visits during pregnancy of women at their first birth was higher (mean=3.26, 95% CI:3.07-3.45) and lower at above their 3rd birth (mean=1.69, 95% CI:1.53–1.85). Women aged below 20 years at birth attended the highest mean ANC visits (mean=2.88, 95% CI:2.65–3.11) while the lowest value was for women aged above 35 years at birth (mean=2.67, 95% CI:1.58–3.77).

The mean number of ANC visits of women exposed to media was greater (mean=3.48, 95% CI:3.32–3.63), whereas it was lower for women who were not (mean=1.85, 95% CI:1.70–2.00). Mean number of ANC visits during pregnancy was found to the lowest (1.48) (mean=1.48, 95% CI:1.32–1.63) among illiterate women and this figure was the highest for higher educated women (mean=4.88, 95% CI:4.44–5.32). The average ANC visits among women increased with their increasing levels of education. From Table 2, it is also evident that women belonging

Table 2: Frequency (%) distribution, mean ANC visits and 95% confidence interval (CI) for mean along with p-value of ANOVA F-test for the demographic and socio-economic variables.

Variables	Number of women (%)	Mean ANC visits	95% CI	p-value
Administrative division				0.012
Barisal	521 (11.9)	2.45	(2.24, 2.65)	
Chittagong	842 (19.2)	2.72	(2.44, 3.00)	
Dhaka	777 (17.7)	3.15	(2.85, 3.45)	
Khulna	515 (11.7)	3.64	(3.21, 4.07)	
Rajshahi	535 (12.2)	3.00	(2.58, 3.42)	
Rangpur	542 (12.3)	3.21	(3.00, 3.42)	
Sylhet	658 (15.0)	2.03	(1.86, 2.21)	
Place of residence				< 0.001
Urban	1407 (32.1)	3.82	(3.58, 4.05)	
Rural	2983 (67.9)	2.42	(2.29, 2.54)	
Birth order				< 0.001
1st birth	1778 (40.5)	3.26	(3.07, 3.45)	
2nd to 3rd birth	1998 (45.5)	2.87	(2.70, 3.05)	
Above 3rd birth	614 (14.0)	1.69	(1.53, 1.85)	
Mother's age at birth				0.791
<20	1374 (31.3)	2.88	(2.65, 3.11)	
20 − 35	2835 (64.6)	2.87	(2.75, 2.98)	
> 35	181 (4.1)	2.67	(1.58, 3.77)	
Exposure of media				< 0.001
Non-exposure	1658 (37.8)	1.85	(1.70, 2.00)	
Exposure	2732 (62.2)	3.48	(3.32, 3.63)	
Education				< 0.001
No education	583 (13.3)	1.48	(1.32, 1.63)	
Primary	1198 (27.3)	2.09	(1.96, 2.21)	
Secondary	2100 (47.8)	3.20	(3.02, 3.39)	
Higher	509 (11.6)	4.88	(4.44, 5.32)	
Wealth index				< 0.001
Poor	1748 (39.8)	1.95	(1.77, 2.14)	
Middle	840 (19.1)	2.50	(2.35, 2.65)	
Rich	1802 (41.1)	3.92	(3.73, 4.11)	
NGO membership				0.289
Non-member	3432 (78.2)	2.90	(2.77, 3.02)	
Member	958 (21.8)	2.75	(2.50, 3.00)	
Women empowerment				0.002
No	865 (19.7)	2.51	(2.34, 2.67)	
Yes	3525 (80.3)	2.95	(2.82, 3.09)	
Education gap				0.308
Husband more	777 (17.7)	2.74	(2.57, 2.91)	
Wife more or equal	3613 (82.3)	2.89	(2.76, 3.02)	
Religion				0.266
Islam	4036 (91.9)	2.84	(2.72, 2.97)	
Others	354 (8.1)	3.08	(2.80, 3.36)	
Mothers having son				0.044
No son	1415 (32.2)	3.03	(2.85, 3.22)	
Else	2975 (67.8)	2.78	(2.64, 2.93)	
BMI				< 0.001
Normal weight	2573 (58.6)	2.83	(2.67, 2.99)	
Under weight	1077 (24.5)	2.20	(1.98, 2.42)	
Over weight	620 (14.1)	3.80	(3.57, 4.02)	
Obesity	120 (2.8)	4.73	(4.20, 5.25)	

to a rich family received the highest (mean=3.92, 95% CI:3.73–4.11) average ANC utilization and it was also the lowest (mean=1.95, 95% CI:1.77–2.14) for the women belonging to a

poor family. Women who were involved with taking important decisions attended higher (mean=2.95, 95% CI:2.82–3.09) mean ANC visits and it was comparatively lower (mean=2.51, 95% CI:2.34–2.67) for women who were not. In Bangladesh, it is seen that women who having no son took more ANC visits (mean=3.03, 95% CI:2.85–3.22) than mothers having son (mean=2.78, 95% CI:2.64–2.93). Women with under weight attended the lowest (mean=2.20, 95% CI:1.98–2.42) average frequency of ANC visits and this value was the highest (mean=4.73, 95% CI:4.20–5.25) for those women who have the obesity.

For modelling and multivariate analysis of ANC count data, we recall that the variance of response variable (number of ANC visits) was greater than mean (Table 1). It follows immediately that the overdispersion should be taken under consideration in the data analysis. This was done by detecting the overdispersion first and then testing whether it is statistically significant or not. In order to detect the presence of overdispersion, we first fitted PR model and computed the value of dispersion by using Pearson χ^2 statistic. This value was 4.36 (greater than 1) which clearly indicates the existence of overdispersion in ANC count data and can be found in Table 3. From *p*-values of both score ($p = 0.008$) and Lagrange multiplier ($p < 0.001$) tests in the right panel of Table 3, it is evident that the presence of overdispersion was statistically highly significant and the fitted PR model is overdispersed.

Table 3: Detection of overdispersion using Pearson residual χ^2 statistic and tests of significance of Poisson overdispersion along with AIC values for different models.

Detection and model Selection			Tests of overdispersion (PR)	
Models	AIC	*Dispersion*	Methods	*p*-value
PR	20716	4.36	Dean and Lawless	0.008
Q_PR	20716	4.36	Winkelmann	0.008
NBR	18317	1.97	Cameron and Trivedi	0.008
MPR_ILRE	**18049**	**0.49**	Lagrange multiplier	<0.001

We now compared performances of different regression models in order to select the best one for analyzing the overdispersed ANC attendance count data of women in Bangladesh. The AIC and *Dispersion* values were computed for different models: PR, Q_PR, NBR and MPR_ILRE and given in the left panel of Table 3. It is seen that AIC value (18049) is minimum and *Dispersion* value (0.49) is also close to 1 for MPR_ILRE compared with other models. It follows that MPR_ILRE is the best choice for analyzing ANC visits count data because of its smallest AIC value and the overdispersion has also been captured very well by MPR_ILRE. Finally, determinants of ANC visits among women during their pregnancy in Bangladesh have been identified by fitting the MPR_ILRE model and the results are summarized in Table 4.

From IRR and *p*-values in Table 4, it is clearly seen that women living in Khulna ($p <0.001$, IRR=1.19, 95%CI:1.051–1.321) and Rangpur ($p <0.001$, IRR=1.25, 95%CI: 1.101–1.401) divisional cities were 1.19 and 1.25 times as likely to have higher average ANC visits during their pregnancy compared to women who live in Dhaka region and these regions were found to be statistically significant for ANC attendance of pregnant women in Bangladesh. The average ANC visits of women living in Chittagong ($p=0.002$, IRR=0.87, 95%CI:0.808–0.940) and Sylhet ($p <0.001$, IRR=0.83, 95%CI:0.761–0.895) were 13% and 17% lower, respectively than women from Dhaka division and these were also found statistically significant. The average ANC visits of women during their pregnancy varied significantly in rural and urban places of residence ($p <0.001$, IRR=1.19, 95%CI:1.100–1.270). The urban women were 1.19 times as likely to have ANC visits than women from rural areas. The birth order of women was also found to be significant and we found that the expected number of ANC visits was 7% higher for women at their first birth ($p <0.025$, IRR=1.07, 95%CI:1.101–1.141) and 16% lower for mothers at their fourth and above birth ($p <0.001$, IRR=0.84, 95%CI:0.774–0.910) compared with the women at their 2nd and 3rd birth. The mean number of ANC visits of women having

Table 4:   Results obtained from fitting mixed Poisson regression with individual–level random effects (MPR_ILRE) to overdispersed ANC visits count data of women in Bangladesh.

Variable	Estimate	SE	p–value	IRR	95% CI for IRR
Intercept	0.128	0.079	0.107	1.14	(0.937, 1.337)
Administrative division					
Barisal	-0.070	0.052	0.179	0.93	(0.843, 1.021)
Chittagong	-0.135	0.044	0.002	0.87	(0.808, 0.940)
Khulna	0.171	0.049	<0.001	1.19	(1.051, 1.321)
Dhaka (ref.)	–	–	–	–	–
Rajshahi	-0.040	0.050	0.422	0.96	(0.871, 1.051)
Rangpur	0.224	0.049	<0.001	1.25	(1.101, 1.401)
Sylhet	-0.189	0.050	<0.001	0.83	(0.761, 0.895)
Place of residence					
Rural (ref.)	–	–	–	–	–
Urban	0.170	0.031	<0.001	1.19	(1.100, 1.270)
Birth order					
1st birth	0.069	0.031	0.025	1.07	(1.001, 1.141)
2nd and 3rd birth (ref.)	–	–	–	–	–
Above 3rd birth	-0.172	0.049	<0.001	0.84	(0.774, 0.910)
Exposure of media					
Non-exposure (ref.)	–	–	–	–	–
exposure	0.213	0.037	<0.001	1.24	(1.126, 1.348)
Education					
No education (ref.)	–	–	–	–	–
Primary	0.215	0.054	<0.001	1.24	(1.077, 1.403)
Secondary	0.444	0.053	<0.001	1.56	(1.307, 1.811)
Higher	0.659	0.064	<0.001	1.93	(1.464, 2.402)
Wealth index					
Poor	-0.091	0.042	0.031	0.91	(0.844, 0.982)
middle (ref.)	–	–	–	–	–
Rich	0.235	0.040	<0.001	1.27	(1.140, 1.390)
Women empowerment					
No (ref.)	–	–	–	–	–
Yes	0.070	0.035	0.045	1.07	(0.994 1.152)
Mother having son					
No son (ref.)	–	–	–	–	–
Else	-0.014	0.030	0.631	0.97	(0.929, 1.043)
BMI					
Normal weight (ref.)	–	–	–	–	–
Under weight	-0.126	0.035	<0.001	0.88	(0.829, 0.935)
Over weight	0.165	0.039	<0.001	1.18	(1.073, 1.285)
Obesity	0.296	0.075	<0.001	1.34	(1.078, 1.610)

ref. indicates reference group

access to media was 24% (p <0.001, IRR=1.24, 95%CI:1.126–1.348) was higher than women who do not have any access to mass-media. There is a significant positive impact of exposure of media on ANC visits of women during their pregnancy period in Bangladesh.

Women with primary education (p <0.001, IRR=1.24, 95%CI:1.077–1.403), secondary education (p <0.001, IRR=1.56, 95%CI:1.307–1.811) and higher education (p <0.001, IRR=1.93, 95%CI:1.464–2.402) attended 24%, 56% and 93% higher average ANC visits, respectively during their pregnancy than women with no education. It is also found that education has highly significant effect on ANC attendance. On average, the frequency of ANC visits was 9% less for women in poor family (p=0.031, IRR=0.91, 95%CI:0.844–0.982) and 27% (IRR=1.260) higher for women belonging to rich family (p <0.001, IRR=1.27, 95%CI:1.140–1.390) than women in middle class family; and wealth index showed the statistical significant impact on ANC

visits of women. Women who were empowered (p=0.045, IRR=1.07, 95%CI:0.994–1.152) and having son (p=0.631, IRR=0.97, 95%CI:0.929–1.043) received only 7% higher and 3% lower ANC visits on average, respectively. The findings revealed that the variables: women empowerment and mothers having son were statistically insignificant for average ANC visits. The average ANC utilization for women with under weight (p <0.001, IRR=0.88, 95%CI:0.829–0.935) was 12% lower whereas the corresponding values were 18% and 12% higher for women with overweight (p <0.001, IRR=1.18, 95%CI:1.073–1.285) and obesity (p <0.001, IRR=1.34, 95%CI:1.078–1.610), respectively than normal weighted women. The BMI of women showed the significant effect on receiving ANC attendance during their pregnancy in Bangladesh.

4. Conclusion

In this study, we used data obtained from the latest BDSH 2014 survey for estimating the current scenario of ANC utilization and its potential determinants of women during their pregnancy in Bangladesh. It was estimated that only 6.5% women received the WHO (2016) recommended eight or higher ANC visits during their pregnancy. It was also found that women received on average less than three ANC visits during their whole course of pregnancy. The variance of number of ANC visits was found to be higher than mean and hence the equidispersion assumption of Poisson regression was violated. We detected and tested the amount of overdispersion present in the data and found to be statistically significant. Thus, we suggest that one should account for the consequence of overdispersion giving high importance to assess the estimates of regression parameters precisely in the count data analysis. Study findings reveal that the performance of MPR_ILRE model was comparatively higher than other three GLMs: PR, Q_PR and NBR in the presence of substantial overdispersion to ANC data of women in Bangladesh. Thus, the significant factors affecting adequate ANC visits of women were determined by using the selected MPR_ILRE model in multivariate analysis.

The findings of this study showed that administrative division or region, place of residence, birth order, exposure of mass-media, education, wealth index and BMI of Bangladeshi women have significant effect on ANC visits during their pregnancy. Women living in urban areas, belonging to rich family, with higher educational attainment, and having mass-media exposure received more ANC visits for the safe pregnancy. According to the findings of this study, significant steps should be taken to aware and motivate women for taking higher ANC visits during the pregnancy period irrespective of their birth orders. Necessary efforts should also be taken to educate women in Bangladesh, to ensure easy access to mass–media and to take care of their weight, should be facilitated for better antenatal care in Bangladesh. More precisely, the concerned authorities need to provide appropriate health care facilities in rural areas and women who belonging to poor families for reducing pregnancy complications and hence to reduce maternal and child mortality in Bangladesh.

Acknowledgements

Data availability statement

We used secondary data from the Demographic and Health Surveys (DHS) Program. The data are available online at: https://dhsprogram.com/data/available-datasets.cfm.

Disclosure statement

The authors declare that they have no conflict of interest.

References

Akaike H (1998). *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer, New York. URL https://link.springer.com/chapter/10.1007/978-1-4612-1694-0_15.

Becker S, Peters DH, Gray RH, Gultiano C, Black RE (1993). "The Determinants of Use of Maternal and Child Health Services in Metro Cebu, the Philippines." *Health Transitional Review*, **3**(1), 77–91. URL https://www.ncbi.nlm.nih.gov/pubmed/10148799.

Bhowmik KR, Das S, Islam MA (2020). "Modelling the Number of Antenatal Care Visits in Bangladesh to Determine the Risk Factors for Reduced Antenatal Care Attendance." *PloS ONE*, **15**(1), e0228215. URL https://doi.org/10.1371/journal.pone.0228215.

Cameron AC, Trivedi PK (1990). "Regression–Based Tests for Overdispersion in the Poisson Model." *Journal of Econometrics*, **46**(3), 347–364. URL https://s100.copyright.com/AppDispatchServlet?publisherName=ELS&contentID=030440769090014K&orderBeanReset=true.

Chakraborty N, Islam MA, Chowdhury RI, Bari W, Akhter HH (2003). "Determinants of the Use of Maternal Health Services in Rural Bangladesh." *Health promotion international*, **18**(4), 327–337. URL https://doi.org/10.1093/heapro/dag414.

Chowdhury RI, Islam MA, Chakraborty N, Akhter HH (2007). "Determinants of Antenatal Morbidity: A Multivariate Analysis." *World health & population*, **9**(3), 9–18. URL https://doi.org/10.12927/whp.2007.19038.

Dean C, Lawless JF (1989). "Tests for Detecting Overdispersion in Poisson Regression Models." *Journal of the American Statistical Association*, **84**(406), 467–472. URL https://www.jstor.org/stable/2289931?seq=1#metadata_info_tab_contents.

Dobson AJ, Barnett AG (2018). *An Introduction to Generalized Linear Models*. 4th edition. CRC press. URL https://books.google.com.bd/books/about/An_Introduction_to_Generalized_Linear_Mo.html?id=YOFstgEACAAJ&redir_esc=y.

Harrison XA (2014). "Using Observation–Level Random Effects to Model Overdispersion in Count Data in Ecology and Evolution." *PeerJ*, **2**, e616.

Hilbe JM (2011). *Negative Binomial Regression*. 2nd edition. Cambridge University Press. URL https://link.springer.com/article/10.1007/s11336-012-9263-7.

Hossain Z, Akter R, Sultana N, Kabir E (2020). "Modelling Zero-Truncated Overdispersed Antenatal Health Care Count Data of Women in Bangladesh." *PloS ONE*, **15**(1), e0227824. URL https://doi.org/10.1371/journal.pone.0227824.

Islam M, Sultana N (2019). "Risk Factors for Pregnancy Related Complications Among Urban Slum and Non-Slum Women in Bangladesh." *BMC pregnancy and childbirth*, **19**(1), 235. URL https://doi.org/10.1186/s12884-019-2392-6.

Islam MM, Masud MS (2018a). "Determinants of Frequency and Contents of Antenatal Care Visits in Bangladesh: Assessing the Extent of Compliance with the WHO Recommendations." *PloS ONE*, **13**(9). URL https://doi.org/10.1371/journal.pone.0204752.

Islam MM, Masud MS (2018b). "Health Care Seeking Behaviour During Pregnancy, Delivery and the Postnatal Period in Bangladesh: Assessing the Compliance with WHO Recommendations." *Midwifery*, **63**, 8–16. URL https://doi.org/10.1016/j.midw.2018.04.021.

Jayaraman A, Chandrasekhar S, Gebreselassie T (2008). "Factors Affecting Maternal Health Care Seeking Behavior in Rwanda." URL https://www.dhsprogram.com/pubs/pdf/WP59/WP59.pdf.

Latif AM, Hossain MZ, Islam MA (2008). "Model Selection Using Modified Akaike's Information Criterion: An Application to Maternal Morbidity Data." *Austrian Journal of Statistics*, **37**(2), 175–184. URL https://doi.org/10.17713/ajs.v37i2.298.

Magnusson A, Skaug H, Nielsen A, Berg C, Kristensen K, Maechler M, Bentham KV, Koen N, Sadat B, Bolker B, Brooks M (2020). "Package'glmmTMB': Generalized Linear Mixed Models using Template Model Builder." *The Comprehensive R Archive Network*. URL https://cran.r-project.org/web/packages/glmmTMB/glmmTMB.pdf.

McCullagh P, Nelder JA (1989). *Generalized Linear Models*. 2nd edition. Chapman & Hall, London.

Mohammad KA, Zahura FT, Rahman MM (2017). "Importance of Maternal Education on Antenatal Care Visits in Bangladesh." *Bangladesh Journal of Scientific Research*, **30**(1-2), 23–33.

Navaneetham K, Dharmalingam A (2002). "Utilization of Maternal Health Care Services in Southern India." *Social Science & Medicine*, **55**, 1849–1869. URL https://doi.org/10.1016/S0277-9536(01)00313-6.

NIPORT, ICDDRB, MEASURE Evaluation (2017). "Bangladesh Maternal Mortality and Health Care Survey 2016: Preliminary Report." URL https://www.measureevaluation.org/resources/publications/tr-17-218.

NIPORT, Mitra & Associates, ICF International (2016). "Bangladesh Demographic and Health Survey 2014." URL https://dhsprogram.com/pubs/pdf/FR311/FR311.pdf.

Nisar N, White F (2003). "Factors Affecting Utilization of Antenatal Care Among Reproductive Age Group Women (15-49 years) in An Urban Squatter Settlement of Karachi." *Journal of Pakistan Medical Association*, **53**(2), 47. URL https://ecommons.aku.edu/pakistan_fhs_mc_chs_chs/459.

Pandit RD (1992). "Role of Antenatal Care in Reducing Maternal Mortality." *Asia-Oceania Journal of Obstetrics and Gynaecology*, **18**(1), 1–6. URL https://doi.org/10.1111/j.1447-0756.1992.tb00291.x.

Rahman KM (2009). "Determinants of Maternal Health Care Utilization in Bangladesh." *Research Journal of Applied Sciences*, **4**(3), 113–119. URL http://medwelljournals.com/abstract/?doi=rjasci.2009.113.119.

Rahman MM, Hossain MA (2019). "Impact of Community Education on Antenatal Care Visits in Bangladesh: A Multilevel Analysis." *The Dhaka University Journal of Science*, **67**(1), 41–46. URL https://www.researchgate.net/publication/332651718_Impact_of_Community_Education_on_Antenatal_Care_Visits_in_Bangladesh_A_Multilevel_Analysis.

Stroup WW (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. CRC press. URL https://books.google.com.bd/books/about/Generalized_Linear_Mixed_Models.html?id=GcGrySpkXRMC&redir_esc=y.

Sultana N, Bari W (2017). "Analyzing Overdispersed Antenatal Care Visits of Pregnant Women in Bangladesh: Negative Binomial Regression Model." *Dhaka University Journal of Science*, **65**(2), 133–137. URL https://www.researchgate.net/publication/332875205_Analyzing_overdispersed_antenatal_care_visits_of_pregnant_women_in_Bangladesh_Negative_binomial_regression_model.

Titaley CR, Hunter CL, Heywood P, Dibley MJ (2010). "Why Don't Some Women Attend Antenatal and Postnatal Care Services?: A Qualitative Study of Community Members' Perspectives in Garut, Sukabumi and Ciamis Districts of West Java Province, Indonesia." *BMC pregnancy and childbirth*, **10**(1), 61. URL http://www.biomedcentral.com/1471-2393/10/61.

Ver Hoef JM, Boveng PL (2007). "Quasi-Poisson vs. Negative Binomial Regression: How Should We Model Overdispersed Count Data?" *Ecology*, **88**(11), 2766–2772. URL https://www.jstor.org/stable/27651434?seq=1#metadata_info_tab_contents.

WHO (2007). "Standards for Maternal and Neonatal Care." URL https://www.who.int/reproductivehealth/publications/maternal_perinatal_health/a91272/en/.

WHO (2016). "WHO Recommendations on Antenatal Care for a Positive Pregnancy Experience." URL http://apps.who.int/iris/bitstream/10665/250796/1/9789241549912-eng.pdf?ua=1.

WHO (2018). "SDG 3: Ensure Healthy Lives and Promote Wellbeing for All At All Ages." URL https://www.who.int/sdg/targets/en/.

WHO (2019). "Maternal Mortality." URL https://www.who.int/news-room/fact-sheets/detail/maternal-mortality.

Winkelmann R (2008). *Econometric Analysis of Count Data*. 5th edition. Springer, New York. URL https://www.springer.com/gp/book/9783540776482.

Affiliation:

Zakir Hossain (corresponding author)
Department of Statistics
University of Dhaka
Dhaka 1000, Bangladesh
E-mail: zakir.hossain@du.ac.bd
URL: https://www.du.ac.bd/faculty/faculty_details/STA/2329

Maria
Department of Statistics
University of Dhaka
Dhaka 1000, Bangladesh
E-mail: maria.mukta11@gmail.com