



Modeling Determinants of Low Birth Weight for Under Five-Children in Ethiopia

Daniel Biftu Bekalo¹ · Zelalem Mehari Nigussie² · Belay Birlie Yimer³ · Zelalem Alamrew Anteneh²

Received: 8 January 2019 / Revised: 17 February 2020 / Accepted: 1 May 2020 / Published online: 28 May 2020
© The Author(s) 2020

Abstract

Low birth weight (LBW) is a major determinant of morbidity, mortality and disability in infancy and childhood and has a long-term impact on health outcomes in adult life. This study was aimed to model LBW using marginal and generalized linear mixed models as well as identify the potential risk factors of LBW in Ethiopia. Data was taken from the 2011 Ethiopian demographic and health survey, which is a nationally representative survey of children in the 0–59 month age groups. Two model families, generalized estimating equation and alternating logistic regression models from marginal model family, and generalized linear mixed model from cluster specific model family were used for the analysis. The result showed that 34.8% of children were born with LBW. Alternating logistic regression model was best fits the data for population-averaged effects of the given factors on birth weight than generalized estimating equation model. Generalized linear mixed model with two random intercepts was the best model to evaluate within and between regional heterogeneity of birth weight. Both the best-fitted models gave the same conclusion that sex, wealth status, age, antenatal care, marital status, vaccination, anemia and mother education level were the determinant factors of LBW. This study contributes to the understanding of the individual and collective effect of maternal, socio-economic and child related factors influencing infant birth weight in Ethiopia.

Keywords Low birth weight · Generalized estimating equation · Alternating logistic regression · Generalized linear mixed model

Abbreviations

LBW	Low birth weight
CLID	Cluster level ID
AIC	Akaike information criterion
BIC	Bayesian information criterion

✉ Daniel Biftu Bekalo
danibiftu@gmail.com

Extended author information available on the last page of the article

QIC	Quasi information criterion
SE	Standard error
CSA	Central Statistical Agency
EDHS	Ethiopia Demographic and Health Survey
WHO	World Health Organization
GEE	Generalized estimating equation
ALR	Altering logistic regression
GLMM	Generalized linear mixed model

1 Background

One of the poor outcomes of pregnancy that has caught the attention of the World Health Organization (WHO) is LBW. LBW has been defined as weight at birth of less than 2500 g [1]. The incidence of LBW is estimated to be 16% worldwide, 19% in the least developed and developing countries and 7% in the developed countries [1]. Globally, more than 20 million infants are born with LBW [1]. Low birth weight (LBW) can be caused either by premature delivery (short gestation < 37 week) or by foetal growth retardation. Evidences showed that low maternal food intake, hard physical work during pregnancy, and illness, especially infections are known factors for pre-term delivery and foetal growth retardation(source). What is more, cigarette smoking, genetic and environmental factors, short maternal stature, very young age, high parity, close birth spacing are associated with LBW [2].

Low birth weight is a worldwide concern, with LBW newborns accounting for 15.5% of all births [3]. This concern exists in both developed and developing countries; however, the burden is more pronounced in developing countries, with 95.6% of all LBW births occurring in these countries [3]. The region of the world with the highest occurrence of LBW newborns is South-central Asia, where 27.1% of infants are born with a LBW. The regions with the next highest proportions of LBW newborns are Western Africa and Western Asia (both 15.4%) [3]. The prevalence of LBW in sub-Saharan Africa ranges between 13 and 15%, with little variation across the region as a whole. In East Africa the prevalence of LBW is 13.5% [1].

LBW is one of the critical issues in Ethiopia that causes short-term and long-term health consequences among babies, and tends to have higher mortality and morbidity. In Ethiopia, the magnitude of LBW babies has increased in the past 5 years, it ranged from 8% in 2000 to 14% in 2011 [4].

LBW is a reasonable well-defined problem caused by factors that are potentially modifiable and the costs of preventing them are well within reach, even in poor countries like Ethiopia. Therefore, it is very important to determine the factors of LBW in various communities in the country in order to come up with feasible intervention strategies to minimize the problem.

This study aims at finding the magnitude and the determinants of low birth weight in Ethiopia based on the 2011 EDHS data by taking into consideration various maternal, socio-economic, demographic and environmental factors. Moreover, previous studies on this area in Ethiopia were considered about modeling only the fixed effects of covariates without including the random effects and with no considering

sampling structures of data. Most of the studies previously done are simply using only the ordinary logistic regression model.

Thus, the little magnitude of this service and lack of appropriateness of the model applied for clustered data have generated interest in assessing determinant factors affecting low birth weight by fitting a statistical model that can explain the data in most meaningful manner. This study, therefore, has tried to fill the gaps in understanding the status of child weight at birth by identifying determinant factors of LBW in Ethiopia and assessing the performance of different models using clustered data from EDHS 2011 by addressing the following research questions:

- Which covariates are the most determinant factors for LBW?
- Which fitted model for the birth weight is statistically plausible?
- Is there a significant within and between regional heterogeneity of weights of child at birth?

2 Methods

2.1 Data Source

The source of data for this study was the 2011 Ethiopia Demographic and Health Survey (EDHS), which is obtained from Central Statistical Agency (CSA). It was the third survey conducted in Ethiopia as part of the worldwide Demographic and Health Surveys project. The 2011 Ethiopian Demographic and Health Survey, was designed to provide estimates for the health and demographic variables of interest for the following domains. Ethiopia as a whole; urban and rural areas (each as a separate domain); and 11 geographic administrative regions (9 regions and 2 city administrations).

2.2 Study Population

The 2007 Population and Housing Census, conducted by the CSA, provided the sampling frame from which the 2011 EDHS sample was drawn. Administratively, regions in Ethiopia are divided into zones, and zones, into administrative units called weredas. Each wereda was further subdivided into the lowest administrative unit, called kebele. During the 2007 Census, each kebele was subdivided into census enumeration areas (EAs) or clusters, which were convenient for the implementation of the census. The 2011 EDHS sample was selected using a stratified, two-stage cluster sampling design. Clusters were the sampling units for the first stage. The sample included 624 clusters, 187 in urban areas and 437 in rural areas. Households comprised the second stage of sampling. In the second stage, a fixed number of 30 households were selected for each cluster. A complete listing of households was carried out in each of the selected clusters from September 2010 through January 2011 [4].

The 2011 EDHS used three questionnaires: the Household Questionnaire, the Woman's Questionnaire, and the Man's Questionnaire. These questionnaires were adapted from model survey instruments developed for the measure DHS project to reflect the population and health issues relevant to Ethiopia. In addition to English, the questionnaires were translated into three major local languages-Amharigna, Oromiffa, and Tigrigna.

A representative sample of 17,817 households was selected for the 2011 EDHS. A total of 11,654 children (0–59 months) were surveyed. The 2011 EDHS questionnaire recorded birth weight, if available from written records or mother's recall, for all births in the 5 years preceding the survey. Because birth weight may not be known for few babies, and particularly for babies delivered at home and not weighed at birth, the mother's estimate of the baby's weight at birth was also obtained. Although subjective, mothers' estimates can be a useful proxy for the weight of the child. A total of 11,654 children less than 59 months were identified in the households of selected clusters. There were cases in which information on the relevant variables was missing and these cases were excluded from the analysis. Thus, the analysis presented in this study on the risk factors of LBW was based on the 3225 children aged less than 59 months.

2.3 Variables of the Study

2.3.1 Response Variable

The child weight was first dichotomized based on the cut-off points as described in literature review leading to the binary response (Table 1).

2.3.2 Predictor (Explanatory Variables)

The variables that were considered in the research and expected to be the risk factors of LBW, were grouped in to maternal, socio-economic, demographic, and health and environmental factors (Table 2).

2.4 Method of Data Analysis

To address the objectives of the study and to reach remarkable conclusion, descriptive analysis, marginal models, cluster specific models analysis were used. Generalized Estimating Equations (GEE) and Alternating Logistic Regression (ALR) was used to

Table 1 Coding and explanation of response variable

Variable	Presentation of variable	Factor coding
Child weight at birth	Child weight	1 = Small birth weight (< 2500 g) 0 = Large birth weight (\leq 2500 g)

Table 2 Coding and explanation of explanatory variables

Covariates	Categories
Sex	0 = female, 1 = male
Wealth status	1 = poor, 2 = middle, 3 = rich
Residence	0 = rural, 1 = urban
Age	1 = 15–19, 2 = 20–39, 3 = 40–49
Terminated pregnancy	0 = no, 1 = yes
Antenatal visit	1 = no visit, 2 = 1–4, 3 = ≥ 5
Marital Status	1 = married, 2 = widowed, 3 = divorced
Vaccination	0 = no, 1 = yes
Anemia	1 = not anemic, 2 = moderate, 3 = severe
Mother education	1 = no education, 2 = primary, 3 = secondary and above
Birth order	1 = 1–4, 2 = 5–9, 3 = ≥ 10
Preceding birth interval	1 = 1–5, 2 = 6–10, 3 = ≥ 11

analyze the population-averaged effects. Generalized Linear Mixed Model (GLMM) was used to estimate random effects as well as fixed effects in the linear predictors.

2.4.1 Marginal Models

A range of techniques has been developed for analyzing data with categorical response variables. Marginal models are among the statistical models widely used to model clustered or repeated data. The primary objective of marginal model is to analyze the population-averaged effects of the given factors. It includes Generalized Estimating Equations (GEE) and Alternating Logistic Regression (ALR).

2.4.2 Generalized Estimating Equations (GEE)

GEE approach is used to account for the correlation between responses of interest for subjects from the same cluster. It is non-likelihood method that captures the association within clusters in terms of marginal correlations [5]. Liang and Zeger [6] proposed GEE for clustered as well as repeated data, which require only the correct specification of correlation structure. The model for GEE based on generalized linear models and working correlation structure is given by:

$$g(\pi_j) = \text{logit}(\pi_j) = X_j' \beta$$

where $g(\pi_j)$ is logit link function, $X_j = (n_j \times p)$ dimensional vector of known covariates, $\beta = (1 \times p)$ dimensional vector of unknown parameter, and $E(Y_j) = \pi_j$ is expected value of the responses Y_j in cluster j which is binomially distributed as $Y_j \sim \text{bin}(n_j, \pi_j)$.

2.4.3 Parameter Estimation for GEE

GEE is not likelihood approach, rather it is quasi likelihood based. The parameter β is estimated by solving estimating equations which consist of the working correlation matrix R_j and matrix with the marginal variances on the main diagonal and zeros elsewhere A_j . The score equation used to estimate the marginal regression parameters while accounting for the correlation structure is given by:

$$S(\beta) = \sum_{i=1}^m \frac{\partial \pi_j}{\partial \beta'} \left[A_j^{-\frac{1}{2}} R_j A_j^{-\frac{1}{2}} \right]^{-1} (Y_j - \pi_j) = 0.$$

2.4.4 Alternating Logistic Regression (ALR)

ALR is an extension of GEE which measures pair wise association of two observations in the same cluster. ALR extends classical GEE in the sense that precision estimates for both the regression parameters β and the association parameters α . Moreover with ALR inferences can be made about pair wise associations between subjects [5]. Let γ_{jkl} be the log odds ratio between outcomes Y_{jk} and Y_{jl} , and Let $\pi_{jk} = P(Y_{jk} = 1)$ and $v_{jkl} = P(Y_{jk} = 1, Y_{jl} = 1)$, then the association of the two responses is defined as [5]:

$$\text{logit}P((Y_{jk} = 1)/(Y_{jl} = y_{jl})) = \gamma_{jkl}y_{jl} + \log\left(\frac{\pi_{jk} - v_{jkl}}{1 - \pi_{jk} - \pi_{jl} + v_{jkl}}\right)$$

Assume $\gamma_{jkl} = \alpha$, then the pairwise log odds ratio α is the regression coefficient in logistic regression of Y_{jk} on Y_{jl} .

2.4.5 Parameter Estimation for ALR

Rather than maximum likelihood, ALR is also quasi likelihood based. Let ξ_j be a vector with elements $\xi_{jkl} = E(Y_{jk}/(Y_{jl} = y_{jl}))$ and let R_j be the vector of residual with elements $R_{jkl} = Y_{jk} - \xi_{jkl}$. Let S_j a vector of diagonal matrix with diagonal element $\xi_{jkl}(1 - \xi_{jkl})$ and let W_j denote the matrix of $\frac{\partial \xi_j}{\partial \alpha}$. Finally, let $A_j = Y_j - \pi_j$, $B_j = \text{cov}(Y_j)$, $C_j = \frac{\partial \pi_j}{\partial \beta}$. Then the alternating logistic regression parameter $\delta = (\beta, \alpha)$ is the simultaneous solution of the following unbiased estimating equations [6].

$$U_\beta = \sum_{j=1}^m C_j' B_j^{-1} A_j = 0$$

$$U_\alpha = \sum_{j=1}^m W_j' S_j^{-1} R_j = 0$$

The above estimating equation are solving for β and α by using Gauss-Seidel procedure algorithm.

2.5 Cluster Specific Models

Cluster specific models are useful when the interest lies in understanding the response of individual characteristics. It differ from the marginal models by inclusion of parameters that are specific to clusters within a population. Consequently, random effects are directly used in modeling the random variation at different levels of the data.

2.5.1 Generalized Linear Mixed Model (GLMM)

Generalized linear mixed models is one parts of cluster specific models. It extends ordinary regression by allowing non-normal responses and a random component with the link function of the mean. Assumed conditionally on q-dimensional random effects b_j to be drawn independently from $N(0, D)$, The outcomes y_{ij} of Y_j are independent with the density of the form

$$f_j(y_{ij}/b_j, \beta, \phi) = \exp\{\phi^{-1}[y_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(y_{ij}, \phi)\}$$

Then the generalized linear mixed model [5]; with logit link is defined as

$$\text{logit}(\pi_{ij}) = X'_{ij}\beta + Z'_{ij}b_j$$

where $j = 1, 2, \dots, m$ $E(Y_{ij}/b_j) = \pi_{ij}$, is the mean response vector conditional on the random effects b_j , for subjects in cluster j and X_{ij} and Z_{ij} are p-dimensional and q-dimensional vectors of known covariate values. β is a p-dimensional vector of unknown fixed regression coefficients, and ϕ is a scale parameter. The random effects b_j are assumed follow a multivariate normal distribution with mean 0 and covariance matrix D .

2.5.2 Parameter Estimation for GLMM

Random-effects models were fitted by maximization of the marginal likelihood, obtained by integrating out the random effects. The Laplace approximation method [5] has been designed to approximate integrals of the form $I = \int e^{Q(b)} db$ where,

$$Q(b) = \phi^{-1} \sum_{i=1}^{n_j} [y_{ij}(x'_{ij}\beta + z'_{ij}b) - \psi(x'_{ij}\beta + z'_{ij}b)] - \frac{1}{2}b'D^{-1}$$

$Q(b)$ is known, unimodal, and bounded function of a q-dimensional variable b .

2.6 Model Building

In order to select the important factors related to LBW, the backward selection procedure was used. To select significant variables, there are different techniques for

proposed models. Since GEE and ALR models are quasi likelihood based one can use modified AIC called QIC for model selection and model selection in GLMM is based on likelihood ratio test and AIC.

3 Results

Table 3 depicts that, a data of 3225 children (0–59 months old) were included in the analysis; 2102 (65.2%) children were born with large weight whereas 1123 (34.8%) were born with small weight. The finding showed that 39.2% and 30.6% of female and male children were found to be LBW respectively. The proportion of LBW was 38.8%, 35.3% and 25.9% among babies from poor, middle wealth status, and rich mothers respectively. The proportion of LBW 73.3% among babies from mothers whose age was less than 19 years and the figure was 36.8% among babies whose mothers' age was over 40 years. Proportion of bearing child with LBW was 37.8%, 30.4% and 24.3% for mothers with no, one to four, and more than five times of antenatal care (ANC) follow up visits respectively. The study also showed that the percentages of LBW babies for mothers who had vaccination and had no vaccination was 31.2% and 46.0% respectively. The percentage of LBW among mothers who had no history of anemia was 28.9%, however, the proportion is 60.7% for mothers with history of anemia during their pregnancy.

3.1 Analysis of Marginal Models

LBW has been analyzed using marginal models including generalized estimating equation and alternating logistic regression models.

3.1.1 Analysis of Generalized Estimating Equations (GEE)

Under the GEE, model building strategy is started by fitting a model containing all possible covariates in the data. This was done by considering two different working correlation assumptions (*exchangeable and independence*). In order to select the important factors related to LBW, the backward elimination procedure was used. The full model for the probability of getting LBW of i th child from j th cluster, (π_{ij}) was fitted as:

Table 3 Descriptive summary of LBW data

Covariates	Child weight (%)		Total
	Large weight (≤ 2500 g)	Small weight (< 2500 g)	
<i>Mother's age at pregnancy</i>			
15–19	36 (26.7)	99 (73.3)	135
20–39	1795 (67.5)	866 (32.5)	2661
40–49	271 (63.2)	158 (36.8)	429
<i>Mother's education</i>			
No education	1567 (62.9)	923 (37.1)	2490
Primary	484 (72.0)	188 (28.0)	67
Secondary and above	51 (81.0)	12 (19.0)	63
<i>Mother's marital status</i>			
Married	1924 (66.6)	963 (33.4)	2887
Widowed	52 (65.0)	28 (35.0)	80
Divorced	126 (48.8)	132 (51.2)	258
<i>Number of ANC visits</i>			
No ANC visit	1331 (62.2)	809 (37.8)	2140
1–4	575 (69.6)	251 (30.4)	826
≤ 5	196 (75.7)	63 (24.3)	259
<i>Ever had vaccination</i>			
Yes	1681 (68.8)	764 (31.2)	2445
No	421 (54.0)	359 (46.0)	780
<i>Ever had terminated pregnancy</i>			
Yes	411 (59.5)	280 (40.5)	691
No	1691 (66.7)	843 (33.3)	2534
<i>Maternal anemia</i>			
Not anemic	1011 (71.1)	411 (28.9)	1422
Moderate	997 (63.7)	567 (36.3)	1564
Sever	94 (39.3)	145 (60.7)	239
<i>Sex of child</i>			
Female	972 (60.9)	625 (39.2)	1597
Male	1130 (69.4)	498 (30.6)	1628
<i>Birth order</i>			
1–4	1071 (66.9)	585 (33.1)	1656
5–9	912 (65.6)	479 (34.4)	1391
≥ 10	119 (64.7)	59 (35.3)	178
<i>Preceding birth interval</i>			
1–5	1835 (65.3)	975 (34.7)	2810
6–10	193 (64.3)	107 (35.7)	300
≥ 11	74 (64.3)	41 (35.7)	115
<i>Wealth Status</i>			
Poor	1127 (61.2)	714 (38.8)	1841
Middle	349 (64.7)	190 (35.3)	539
Rich	626 (74.1)	219 (25.9)	845

Table 3 (continued)

Covariates	Child weight (%)		Total
	Large weight (≤ 2500 g)	Small weight (< 2500 g)	
<i>Residence</i>			
Rural	1875 (64.7)	1025 (35.3)	2900
Urban	227 (69.8)	98 (30.2)	325
Total	2102 (65.2)	1123 (34.8)	3225

$$\begin{aligned}
\text{logit}(\pi_{ij}) = & \beta_0 + \beta_1 \text{Sex}_M + \beta_2 \text{WealthS}_{Mi} + \beta_3 \text{WealthS}_{Ri} + \beta_4 \text{Age}_1 + \beta_5 \text{Age}_2 \\
& + \beta_6 \text{Antenatalcare}_{1+} + \beta_7 \text{Antenatalcare}_{5+} + \beta_8 \text{Maritalst}_W \\
& + \beta_9 \text{Maritalst}_D + \beta_{10} \text{Vaccination}_Y \\
& + \beta_{11} \text{Anemia}_{Mo} + \beta_{12} \text{Anemia}_{Se} + \beta_{13} \text{Educatinle}_{Pr} \\
& + \beta_{14} \text{Educationle}_{Sec} + \beta_{15} \text{Residence}_U \\
& + \beta_{16} \text{termpregnancy}_Y + \beta_{17} \text{Birthorder}_{5-9} + \beta_{18} \text{Birthorder}_{10+} \\
& + \beta_{19} \text{Prebirthinterval}_{5-10} \\
& + \beta_{20} \text{Prebirthinterval}_{11+}.
\end{aligned}$$

The subscripts in each covariate is defined as, M = Male, Mi = middle, Ri = Rich, U = Urban, 1 = 20–39, 2 = 40–49, Y = Yes, 1+ = 1–4, 5+ = five and above, W = Widowed, D = divorced, Mo = Moderate, Se = severe, Pr = Primary, Sec = Secondary, 10+ = ten and above, 11+ = eleven and above.

After fitting the model, covariates with the largest p value are removed and the model was refitted with the rest of the covariates sequentially. Then, residence, ever had terminated pregnancy, birth order and preceding birth interval are the covariates excluded from the model. The QIC values of full model and reduced models are 4011.6165 (which is found in appendix) and 3986.4033 respectively. Then it turned out that the model with sex, wealth status, age of mother, number of antenatal care, marital status, vaccination, anemia level and mothers' education level was the most parsimonious model.

Then, from Table 4, exchangeable working correlation assumption was found to be plausible since the two standard errors were closer to each other with correlation parameter ($\alpha = 0.0857$). Therefore, the final proposed generalized estimating equation model for low birth weight is given as:

$$\begin{aligned}
\text{logit}(\pi_{ij}) = & \beta_0 + \beta_1 \text{Sex}_M + \beta_2 \text{WealthS}_{Mi} + \beta_3 \text{WealthS}_{Ri} + \beta_4 \text{Age}_1 \\
& + \beta_5 \text{Age}_2 + \beta_6 \text{Antenatalcare}_{1+} + \beta_7 \text{Antenatalcare}_{5+} + \beta_8 \text{Maritalst}_W \\
& + \beta_9 \text{Maritalst}_D + \beta_{10} \text{Vaccination}_Y \\
& + \beta_{11} \text{Anemia}_{Mo} + \beta_{12} \text{Anemia}_{Se} + \beta_{13} \text{Educatinle}_{Pr} \\
& + \beta_{14} \text{Educationle}_{Sec}.
\end{aligned}$$

Table 4 Empirical and model based standard errors for two proposed working correlation

Coeff.	Estimates	Exchangeable		Estimates	Independent	
		Model based (SE)	Empirical (SE)		Model based (SE)	Empirical (SE)
β_0	0.6468	0.2481	0.2523	0.7185	0.2300	0.2820
β_1	− 0.3449	0.0756	0.0779	− 0.3536	0.0769	0.0801
β_2	0.0249	0.1064	0.1045	− 0.3536	0.1065	0.1114
β_3	− 0.3505	0.1037	0.1013	− 0.3887	0.0994	0.1039
β_4	− 0.9798	0.2359	0.2532	− 1.0449	0.2190	0.2780
β_5	− 0.8337	0.2542	0.2677	− 0.8565	0.2395	0.2914
β_6	− 0.1361	0.0929	0.0985	− 0.1307	0.0933	0.0996
β_7	− 0.2818	0.1635	0.1538	− 0.2426	0.1633	0.1554
β_8	0.3448	0.1480	0.1674	0.4387	0.1451	0.1908
β_9	0.0198	0.2403	0.2411	0.0698	0.2423	0.2320
β_{10}	− 0.2584	0.0962	0.1020	− 0.2930	0.0931	0.1088
β_{11}	0.2262	0.0811	0.0854	0.2322	0.0814	0.0872
β_{12}	0.6872	0.1671	0.1910	0.8098	0.1620	0.2226
β_{13}	− 0.1951	0.1026	0.1059	− 0.1783	0.1015	0.1111
β_{14}	− 0.3240	0.3351	0.3337	− 0.3595	0.3402	0.3461

Table 5 Empirical and model based standard errors for two proposed working correlation

Covariates	Analysis of GEE parameter estimates empirical standard error estimates						
	Level	Estimate	SE	95% CI	Z	Pr > Z	
Intercept		0.6468	0.2523	0.1523	1.1414	2.56	0.0104*
Sex	Male	− 0.3449	0.0779	− 0.4976	− 0.1923	− 4.43	< 0.0001*
Wealth index	Middle	0.0249	0.1045	− 0.1800	0.2298	0.24	0.8118
	Rich	− 0.3505	0.1013	− 0.5491	− 0.1520	− 3.46	0.0005*
Age of mothers	20–39	− 0.9798	0.2532	− 1.4761	− 0.4835	− 3.87	0.0001*
	40–49	− 0.8337	0.2677	− 1.3585	− 0.3090	− 3.11	0.0018*
Antenatal care	1–4	− 0.1361	0.0985	− 0.3292	0.0571	− 1.38	0.1674
	Five and above	− 0.2818	0.1538	− 0.5831	− 0.1196	− 3.83	0.0069*
Marital status	Divorced	0.3448	0.1674	0.0167	0.6728	2.06	0.0394*
	Widowed	0.0198	0.2411	− 0.4526	0.4923	0.08	0.9345
Vaccination	Yes	− 0.2584	0.1020	− 0.4584	− 0.0584	− 2.53	0.0113*
Anemia	Moderate	0.2262	0.0854	0.0588	0.3936	2.65	0.0081*
	Sever	0.6872	0.1910	0.3129	1.0614	3.60	0.0003*
Education level	Primary	− 0.1951	0.1059	− 0.4027	− 0.1126	− 3.84	0.0056*
	Secondary and above	− 0.3240	0.3337	− 0.9780	0.3299	− 0.97	0.3314
QIC = 3986.4033							
Alpha (α) = 0.0857							

* $p < 0.05$

Parameter estimates and their corresponding empirically corrected standard errors alongside the p values from the final GEE model are presented in Table 5.

3.1.2 Analysis of Alternating Logistic Regression Model (ALR)

Model building for ALR is follows the same procedure in GEE model building strategy. First ALR model is fitted using all proposed covariates. Then the covariate with the large p value is removed. Residence, ever had terminated pregnancy, birth order and preceding birth interval are removed covariates with (p value > 0.05). The QIC values of both saturated and reduced models are 4011.8139 and 3986.1527 respectively. Using the selected covariates and the association parameter α , alternating logistic regression (ALR) model that provides information about pair wise association of observations between two different individuals within the same cluster was fitted. Therefore, the final proposed ALR model included the association parameter for low birth weight is given as:

Table 6 Parameter estimates (empirically corrected standard errors) from ALR

Covariates	Analysis of ALR parameter estimates empirical standard error estimates						
	Level	Estimate	Standard Error	95% C.I		Z	Pr > Z
Intercept		0.6689	0.2510	0.1770	1.1608	2.67	0.0077*
Sex	Male	− 0.3461	0.0778	− 0.4985	− 0.1936	− 4.45	< 0.0001*
Wealth index	Middle	0.0291	0.1044	− 0.1755	0.2337	0.28	0.7805
	Rich	− 0.3522	0.1012	− 0.5505	− 0.1540	− 3.48	0.0005*
Age of mothers	20–39	− 1.0008	0.2520	− 1.4947	− 0.5068	− 3.97	< 0.0001*
	40–49	− 0.8581	0.2670	− 1.3815	− 0.3348	− 3.21	0.0013*
Antenatal care	1–4	− 0.1375	0.0986	− 0.3308	0.0557	− 1.39	0.1630
	Five and above	− 0.2832	0.1537	− 0.5845	− 0.1181	− 1.84	0.0055*
Marital status	Divorced	0.3402	0.1659	0.0152	0.6653	2.05	0.0402*
	Widowed	0.0213	0.2419	− 0.4528	0.4955	0.09	0.9297
Vaccination	Yes	− 0.2582	0.1019	− 0.4580	− 0.0585	− 2.53	0.0113*
Anemia	Moderate	0.2293	0.0853	0.0622	0.3965	2.69	0.0072*
	Sever	0.6874	0.1887	0.3176	1.0573	3.64	0.0003*
Education level	Primary	− 0.1962	0.1056	− 0.4031	− 0.1107	− 1.86	0.0031*
	Secondary and above	− 0.3351	0.3350	− 0.9916	0.3215	− 1.00	0.3172
Alpha (α)		0.4107	0.0879	0.2385	0.5829	4.67	< 0.0001*
QIC = 3986.1527							

* $p < 0.05$

$$\begin{aligned}
\text{logit}(\pi_{ij}) = & \alpha + \beta_0 + \beta_1 \text{Sex}_M + \beta_2 \text{Wealth}_{S_{Mi}} + \beta_3 \text{Wealth}_{S_{Ri}} \\
& + \beta_4 \text{Age}_1 + \beta_5 \text{Age}_2 + \beta_6 \text{Antenatalcare}_{1+} + \beta_7 \text{Antenatalcare}_{5+} \\
& + \beta_8 \text{Maritalst}_W + \beta_9 \text{Maritalst}_D + \beta_{10} \text{Vaccination}_Y + \beta_{11} \text{Anemia}_{Mo} \\
& + \beta_{12} \text{Anemia}_{Se} + \beta_{13} \text{Educatinle}_{Pr} + \beta_{14} \text{Educationle}_{Sec}.
\end{aligned}$$

3.2 Comparison of GEE and ALR Models

Since the likelihood function does not fully specified in marginal models, model comparison is based on quasi likelihood criteria (QIC) which is the modified AIC criteria. From Tables 5 and 6, we found that the QIC values are 3986.4033 and 3986.1527 for the GEE and ALR respectively which is almost equal. However, the empirically corrected standard errors for ALR model are somewhat smaller than their counterpart under the GEE model. This implies that the ALR fits the data with small disturbance than GEE. Moreover, ALR extends beyond classical GEE in the sense that precision estimates follow for both the regression parameters β and the association parameters α . We were also in a position to emphasize that the association is strongly significant ($p < 0.0001$), provided it has been correctly specified, a declaration we could not make in the corresponding exchangeable GEE analysis. Therefore, we can conclude that ALR is the better model for explaining the marginal association between low birth weight and the selected predictor variables. Thus, the interpretation of parameters is based on the final proposed ALR model. Overall, parameter estimates under ALR are slightly less than those of GEE. This difference in parameter estimates from the two models might be due to the fact that ALR takes the associations into account, whereas GEE not consider the association parameter in the model.

3.2.1 Parameter Interpretation of Marginal Models

Table 6 presents parameter estimates and their corresponding empirically corrected standard errors alongside the p values from ALR model. Each parameter β_j reflects the effect of factor X_j on the log odds of the probability of being born with LBW, statistically controlling all the other covariates in the model. Then, the odds ratio of variables is calculated as the exponent of β_j i.e. odds ratio = e^{β_j} . The ALR analysis from Table 6 suggests that, sex of child is significantly related to birth weight of child. After controlling all other variables in the model the odds that a male child born with LBW is $\exp(\beta_1) = \exp(-0.3461) = 0.7074$ (95% CI 0.6074, 0.8239) times lower than the female child. This means the probability that male child born with LBW is 29% lower than that of female. As it has been seen from the result of the ALR model, mothers wealth status is statistically significant on birth weight of child. The estimated odds that child born to a mother who are from highest wealth status is $\exp(-0.3522) = 0.7031$ (95% CI 0.5766, 0.8572) times less likely to have low birth weight compared to the reference group. This implies that the probability of LBW is reduced by 29% for children whose their mother are from highest wealth status when compared with children whose their mothers are from lowest wealth status. In

this study, middle wealth status has no significant effect on LBW of children. There is also a strong association between age of mother and birth weight of child. This implies that, after adjusting all other predictor variables in the model, the estimated odds that child born to a mother who are from age group 20–39 is $\exp(-1.0008) = 0.3675$ (95% CI 0.2242, 0.6024) times lower to have low birth weight compared to reference age group (15–19). This means percentage of low birth weight is decreased by 63% for children whose their mothers are in age group 20–39 when compared to children whose their mothers are in early age group. The estimated odds that child born to a mother who are from age group 40–49 is $\exp(0.8581) = 0.4239$ (95% CI 0.2512, 0.7154) times lower to have low birth weight when compared to reference age group. This means percentage of low birth weight is decreased by 57% for children whose their mothers are in age group 40–49 when compared to children whose their mothers are in early age group. The results also indicate a negative association between LBW and the number of antenatal care visits. The results suggest that the higher the number of antenatal visits, the lower the odds of LBW. The odds that a child born to mother who follow antenatal care for more than five times is $\exp(-0.2832) = 0.7533$ (95% CI 0.5573, 0.8886) times lower to have low birth weight compared to one whose mother do not follow antenatal care. This implies that low birth weight is reduced by 25% for children whose their mothers follow antenatal care for more than five times. As we can see from the analysis, following antenatal care for less than five times has no significant effect on LBW of child. Another significant ingredient of LBW is marital status of mother. Mothers who are divorced are more likely to deliver child with LBW than mothers who are married. The odds of LBW for divorced mother is $\exp(0.3402) = 1.4052$ (95% CI 1.0153, 1.9450) times higher as compared to reference group. This implies LBW of baby increased by 40% for divorced mothers when compared to married mothers. Statistically significant association has been seen between vaccination and LBW of child. The odds that a child born to vaccinated mother is $\exp(-0.2582) = 0.7724$ (95% CI 0.6325, 0.9431) times lower to have low birth weight compared to one whose mother is not vaccinated. This implies LBW is decreased by 22% for children whose their mothers are vaccinated. Statistically significant association has been seen between LBW and anemia level. The odds that a child born to mother who moderately suffered from anemia is $\exp(0.2293) = 1.2577$ (95% CI 1.0641, 1.4866) times higher to have low birth weight. And the odds that a child born to mother who severely suffered from anemia is $\exp(0.6874) = 1.9885$ (95% CI 1.3738, 2.8785) times higher to have low birth weight compared to one whose mother is not suffered from anemia. This implies that the percentage of delivering child with LBW is increased by 26% and 99% respectively for moderately anemic and severely anemic mothers compared to not anemic mothers. The analysis from Table 6 suggests that, education is significantly related to LBW of children. After controlling all other variables in the model, the odds that mother whose her education level is primary deliver a child with LBW is $\exp(-0.1962) = 0.8218$ (95% CI 0.6682, 0.8952) times lower when compared to the reference group. This shows LBW is reduced by 18% for children whose their mothers education level is primary compared to children whose their mothers are not educated. The ALR model also presents the estimated constant log odds ratio (alpha) which, provide information about the association between individual

observations within the same cluster. The estimated pair wise odds ratio relating two responses from the same cluster is $\exp(0.4107) = 1.5078$ (95% CI 1.2693, 1.7912). Thus, the value of alpha which is greater than one indicates that, the associations is found to be significant (p value < 0.0001) and this means that there is a strong positive association between individual children regarding LBW in the same cluster.

3.3 Analysis of Generalized Linear Mixed Model (GLMM)

Under the GLMM, model fitting began by adoption of the marginal model covariates. Additionally, the model also included the random effects in this case, random intercepts to address the between and within-regional variations. First, main effect covariates and the two random intercepts model were fitted and as usual, non-significant covariates were removed sequentially starting from variables with highest p value for fixed effect covariates. The saturated models for GLMM were fitted with b_j and b_{ij} , where b_j and b_{ij} are two random intercepts.

In order to decide on the better of the two random effects models, two models were fitted, one the saturated model with two random intercepts to estimate between and within regional variations and the other with one random intercept model to estimate within regional variation. AIC and Likelihood ratio test (LRT) were used to compare the two models to select an appropriate models.

Where σ_w and σ_B are within and between regional standard deviation respectively. As we have seen from Table 7, the AIC of model with two random intercept is reduced from 3933.1 to 3919.7 and the log likelihood ratio is reduced from 3889.1 to 3873.8. The small p value of the log likelihood ratio test ($p < 0.001$) also indicates that the model with two random intercept is parsimonious model. p is the p value of the log likelihood ratio test of the two models. Also when considered a model without random effects (i.e simply the generalized linear model), it gives AIC value of 3980.1 which is large as compared to the above two models with random effects.

Next, the covariates for the fixed effect were assessed and the candidate covariates were selected by removing covariates starting from with highest p value sequentially. Then the first removable covariate is preceding birth interval with the highest p value 0.8391 and refitted the reduced model with the remaining covariates. The AIC is reduced from 3919.7 to 3916.0 and the p value of log likelihood ratio test ($p = 0.8556$) supports the reduced model is preferable one. The next removable variable is ever had terminated pregnancy with p value ($p = 0.2345$) and refitted the reduced model. The AIC is reduced from 3916.0 to 3915.4 and the p value of log likelihood ratio test ($p = 0.2359$) supports the reduced model is preferable. The

Table 7 Information criteria for comparison of two random models

Models	AIC	BIC	− 2LogLik	σ_w	σ_B	p value
Model with one random intercept	3933.1	4066.8	3889.1	0.6032		
Model with two random intercept	3919.7	4059.5	3873.8	0.5392	0.2411	0.000

next removable variable is birth order with p value ($p = 0.1734$) and refitted the reduced model. The AIC is reduced from 3915.4 to 3914.3 and the p value of log likelihood ration test ($p = 0.2345$) support the reduced model is preferable. The next removable variable is place of residence with p value ($p = 0.1342$) and refitted the reduced model. For this model AIC is similar with the previously reduced model but still the log likelihood ratio test indicates that the reduced model is better with p value ($p = 0.1096$). In addition, the model with small number of covariates is considered to be preferable. Therefore, the final proposed GLMM for low birth weight of children is given as:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_0 + \beta_1 \text{Sex}_M + \beta_2 \text{Wealth}S_{Mi} + \beta_3 \text{Wealth}S_{Ri} \\ & + \beta_4 \text{Age}_1 + \beta_5 \text{Age}_2 + \beta_6 \text{Antenatalcare}_{1+} + \beta_7 \text{Antenatalcare}_{5+} \\ & + \beta_8 \text{Maritalst}_W + \beta_9 \text{Maritalst}_D + \beta_{10} \text{Vaccination}_Y + \beta_{11} \text{Anemia}_{Mo} \\ & + \beta_{12} \text{Anemia}_{Se} + \beta_{13} \text{Educatinle}_{Pr} + \beta_{14} \text{Educationle}_{Sec} + b_i + b_{ij} \end{aligned}$$

where b_i and b_{ij} are regional and cluster level random intercepts respectively.

The parameter estimate and standard error of GLMM are presented in Table 8 of below.

3.3.1 Parameter Interpretation of GLMM

Unlike in the marginal models, (GEE and ALR) where parameters are treated as population averages, in the GLMM analysis, parameter interpretation is based on specific subjects or cluster. The parameter interpretation is conditional on the random effects, which is common for all individual children in the same cluster. Given the same random intercept b_j , the estimated odds of LBW of child is

Table 8 Parameter estimates (standard errors) and corresponding p value for GLMM

Effect	Level	Estimates	SE	95% CI	p value
Intercept		0.7062	0.2757	[0.1657, 1.2468]	0.0104
Sex	Male	− 0.3815	0.0819	[− 0.5420, − 0.2209]	0.0079
Wealth index	Middle	0.0650	0.1165	[− 0.1633, 0.2934]	0.5768
	Rich	− 0.3304	0.1113	[− 0.5485, − 0.1122]	0.0029
Age of mothers	20–39	− 1.1031	0.2502	[− 1.5937, − 0.6126]	0.0070
	40–49	− 0.9378	0.2706	[− 1.4682, − 0.4074]	0.0005
Antenatal care	1–4	− 0.1557	0.1002	[− 0.3522, 0.0406]	0.1201
	≥ 5	− 0.2956	0.1744	[− 0.6376, − 0.0463]	0.0002
Marital status	Divorced	0.3587	0.1591	[− 0.4221, 0.5855]	0.0242
	Widowed	0.0817	0.2570	[0.0467, 0.6707]	0.7506
Vaccination	Yes	− 0.2640	0.1063	[− 0.4724, − 0.0555]	0.0130
Anemia	Moderate	0.2459	0.0885	[0.0725, 0.4194]	0.0054
	Sever	0.7822	0.1820	[0.4255, 1.1390]	0.0016
Education level	Primary	− 0.1577	0.1117	[− 0.3767, − 0.0112]	0.0030
	Secondary and above	− 0.4141	0.3591	[− 1.1180, 0.2897]	0.2488

$\exp(-0.3815) = 0.6828$ (95% CI 0.5815, 0.8017) times lower for male child when compared to female child in the same j th cluster keeping constant the other fixed effect variable in the model. This implies the probability of low birth weight is 32% less likely for male child than female child in the same cluster at the given random effect. In the same way, the estimated odds that a child born to a mother who are from highest wealth status is $\exp(-0.3304) = 0.7186$ (95% CI 0.5778, 0.8939) times lower to have low birth weight compared to the reference group in the same cluster. This shows that the probability of LBW is reduced by 28% for children whose their mother are from highest wealth status when compared with children whose their mothers are from lowest wealth status. The estimated odds that child born to a mother who are from age group 20–39 is $\exp(-1.1031) = 0.3318$ (95% CI 0.2031, 0.5419) times lower to have low birth weight compared to reference age group (15–19). This means percentage of low birth weight is decreased by 67% for children whose their mothers are in age group 20–39 when compared to children whose their mothers are in early age group in the same cluster. The estimated odds that child born to a mother who are from age group 40–49 is $\exp(-0.9378) = 0.3914$ (95% CI 0.2303, 0.6653) times lower to have low birth weight when compared to reference age group. This means percentage of low birth weight is decreased by 61% for children whose their mothers are in age group 40–49 when compared to children whose their mothers are in early age group in the same cluster. At the given constant random effect, The odds that a child born to mother who moderately suffered from anemia is $\exp(0.2459) = 1.2787$ (95% CI 1.0751, 1.5210) times higher to have low birth weight. And the odds that a child born to mother who severely suffered from anemia is $\exp(0.7822) = 2.1862$ (95% CI 1.5303, 3.1236) times higher to have low birth weight compared to one whose mother is not suffered from anemia. This shows

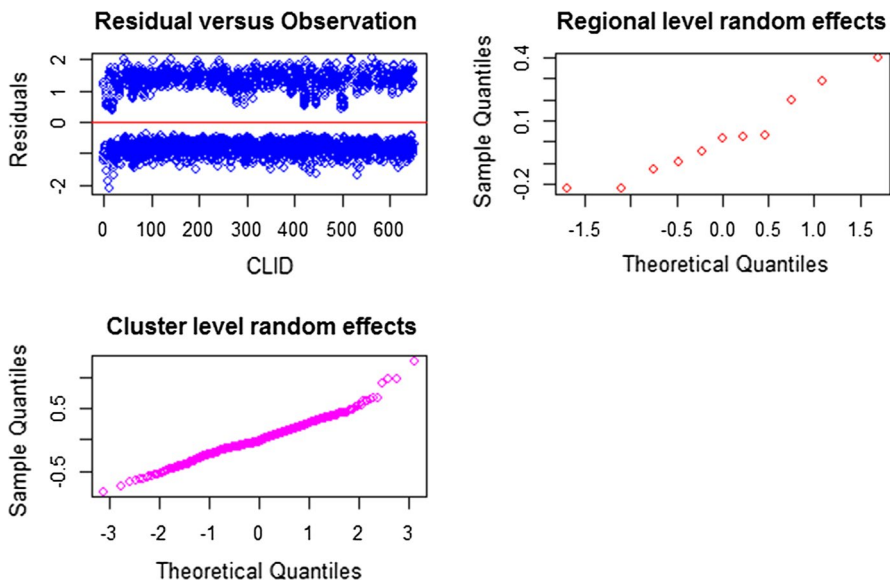


Fig. 1 Diagnosis plots for the generalized linear mixed model

that the probability that mothers deliver child with LBW for mothers who are moderately anemic is 28% more likely than mothers who are not anemic and the probability that mothers deliver child with LBW for severely anemic mothers is two folds more likely than mothers who are not anemic. The interpretation of other predictor variables can be done in a similar manner.

3.3.2 Model Diagnostic for GLMM

Figure 1 below revealed that the residuals versus observation with CLID number suggested that residuals are symmetric around zero (i.e. positive and negative residuals are almost equal). Q–Q plots for normality of random effects at regional and cluster levels illustrates that the random effects are normally distributed with mean zero and variance covariance matrix D. Thus, the fitted GLMM model is fine for LBW data.

4 Discussion

This study was aimed at modeling the determinants of low birth weight in Ethiopia. As a preliminary analysis, assortments of summary statistics were employed to explore the association between the response variable of interest and available covariates. It should be well-known that there is inconsistency in the conclusion from the analysis of various summary statistics, which might be due to the fact that they make use of varying amount of information, which determines the power of their inferences. Thus, the analysis was extended to other statistical methods to account for the clustered nature of correlated observations. The data were then analyzed using two model families one with marginal models (GEE and ALR), and the other is random effects model (Generalized linear mixed model). Two proposed working correlation structures, exchangeable and independence correlation assumptions were taken for the comparison, in GEE model-building strategy.

The model with exchangeable working correlation structure was found to be better fits the data than independence. This supports that considered the clustering nature of the data was essential for the analysis and the dependency of individuals for the given data. In addition, ALR was fitted for simultaneously regress the response variable on explanatory variables as well as association among responses in terms of pair wise odds ratio. Two models from marginal model families were compared in order to assess which model is efficiently explain the relations between response and explanatory variables as well as to evaluate that whether considering pair wise association is important. After then, ALR model was selected as best model and the model shows that there is a positive pair wise association between responses. This is supported the idea explained by Zeger et al., alternating logistic regression is reasonably efficient relative to GEE [6]. The purpose of GLMM was to evaluate within and between regional variations of LBW in Ethiopia.

Two models was fitted one with only one random intercept model to assess only within regional variation and other with two random intercepts model, in order to account within and between regional variations. Additionally, generalized linear model

was fitted as the sake of comparison whether including random effects in the analysis is important or not. The three models were compared using the AIC value followed by likelihood ratio test and we got a model with two random intercept was favorable. This demonstrates that, accounting within and between regional variations for the analysis of LBW should be vital and, indicates within and between regional heterogeneity in LBW. This finding is supported by the explanation or suggestion of Antonio and Beirlant [7]. Even though the two model families are different and their comparability may not be meaningful as they have different parameter interpretations and estimations, parameter estimates obtained from GLMM are generally bigger in absolute values than those from marginal models (GEE and ALR) similar with Agresti [8].

All the fitted models were leads to the same conclusion that sex of child, wealth status, age of mother, number of antenatal care visit, marital status, vaccination, maternal anemia and mother education level were found to be significantly associated with LBW. This study found that male gender has a protective effect against LBW. Male child is less likely to be born with LBW than female child. Which agree with study of Amory et al. [9]. This study finding shows the negative association between wealth status of mothers and LBW which agree with study done in England by Smith et al. [10]. The study shows that the odds of mother bearing child with LBW is consistently decreased as the mother wealth status increased. One of the most predominant causes of low birth weight is the mother's age. The chance of having LBW baby is higher among young mothers of age 15–19. This is similar with finding of Kamaladoss et al. [11].

There was also a significant association between LBW and maternal anemia. According to this study, maternal anemia increased the risk of having a LBW baby. The findings of this study are similar to a study done in Turkey by Chuku [12]. In agreement with previous studies, maternal education emerged as a strong determinant for LBW. Women with 'no education' had the greatest odds of giving birth to an infant with LBW. This finding is similar with some other studies such as, Karim et al. [13]. This study showed the negative effect of number of antenatal care visit on LBW. Those mothers received antenatal care gave birth to higher birth weight babies in comparison to mothers who do not received antenatal care visit.

The other studies also found similar result, Naher et al. [14]. In agreement with previous studies, maternal vaccination emerged as a strong determinant for LBW. Women with 'no vaccination' had the greatest odds of giving birth to an infant with LBW, Som et al. [15]. Another important risk factor for LBW in this study is marital status of mothers. The odds of having infants with low birth weight were higher among mothers who were divorced. However, from the previous studies, residence, terminated pregnancy, birth order and preceding birth interval were significantly associated with LBW; these covariates are not significant determinant factors on this study.

5 Conclusions

For this study two marginal models, GEE and ALR, have been compared for the analysis of marginal or average effects of covariates on the response variable and, we conclude that, ALR model with measure of association exhibited the

best fit for this data than GEE models. For this study also GLMM, with two random intercept model was found to be appropriate for the analysis of within and between regional variations for LBW baby in Ethiopia. This concluded that there is heterogeneity of LBW between and within regions. This study suggests that maternal age, educational level, wealth status, vaccination, child sex and wealth status have negative effect on LBW. Whereas, maternal anemia and marital status have positive effect on LBW. However, in this study, residence, terminated pregnancy, birth order and preceding birth interval were not significantly associated with LBW. More importantly, this study contributes to the understanding of the individual and collective effect of maternal, socio-economic and child related factors influencing infant birth weight in Ethiopia.

This study has identified a number of important factors that influence LBW of baby in Ethiopia. Strategy to reduce LBW in Ethiopia focus has to be given on nutrition education, iron and vitamins supplementation during pregnancy along with discouraging teenage pregnancy. It is suggested that programs that work to reduce the rate of LBW infants should focus on improving maternal lifestyle choices by increasing access, utilization and quality of care, while addressing the intractable socio-economic disparities that continue to indirectly contribute to the incidence of LBW. Socio-economic factors influenced the growth of fetus and outcomes of pregnancies. Most women lacked knowledge of the pregnancy risk factors that adversely affect infant birth weight, and the exact mechanisms by which the risk factors act to cause the adverse effects. Intervention programs and behavior change communication during pregnancy should focus on significant risk factors associated with LBW, and target pregnant women at risk. Health education for pregnant women should be strengthened to promote care seeking and demand for skilled care at all stages of maternity. This way healthy infants are produced who have a better chance of surviving and becoming tomorrow's wealth.

Acknowledgements We acknowledge the Ethiopian Central Statistical Agency for providing us the data.

Author Contributions DBB conceived the idea and drafted the manuscript. DBB analysed and interpreted the data. ZMN was involved in revising the manuscript critically for important intellectual content. BB and ZAA conducted data filtration, edition and cleaning.

Funding No funding was received for this study.

Availability of data and materials We can provide the dataset that has been used during the current study on reasonable request.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Consent for Publication Not applicable.

Ethics Approval and Consent to Participate Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. UNICEF and WHO (2004) Low birth weights: country, regional and global estimates. UNICEF, Editorial and publication section, New York
2. Kramer M (2004) Socioeconomic determinants of intrauterine growth retardation. *Eur Clin Nutr* 52(S1):S21–S28
3. Wardlaw TM (ed) (2004) Low birthweight: country, regional and global estimates. Unicef
4. Central Statistical Agency (2011) Ethiopian demographic and health survey Addis Ababa Ethiopia
5. Molenberghs G, Verbeke G (2005) Models for discrete longitudinal data. Springer
6. Zeger S, Liang K (1986) Longitudinal data analysis using generalized linear models. *Biometrics* 73:13–22
7. Antonio K, Beirlant J (2011) Actuarial statistics with generalized linear mixed models. University Center for Statistics, Leuven
8. Agresti A (2007) An introduction to categorical data analysis, 2nd edn. Wiley, New York
9. Amory JH, Adams KM, Lin MT, Hansen JA, Eschenbach DA, Hitti J (2013) Maternal anemia and prenatal outcome. *Prenat J* 15(3):127–130
10. Smith GC, Smith MF, McNay MB, Fleming JE (2010) First-trimester growth and the risk of low birth weight. *N Engl J Med* 339(25):1817–1822
11. Kamaladoss T, Abel R, Sampathkumar V (2013) Epidemiological correlates of low birth weight in rural Tamil Nadu. *Indian J Pediatr* 59:209–304
12. Chuku SN (2013) Low birth weight in Nigeria: does antenatal care matter? *J Arts Dev Stud* 12:66–71
13. Karim E, Mascie-Taylor CG (1997) The association between birth weight, socio-demographic variables and maternal anthropometry in an urban sample from Dhaka, Bangladesh. *Ann Hum Biol* 24:387–401
14. Naher N, Afroza S, Hossain M (2010) Incidence of LBW in three selected communities of Bangladesh. *Bangladesh Med Res Counc Bull* 24(2):49–54
15. Som S Jr, Pal M, Adak DK, Gharami AK, Bharati S, Bharati P (2012) Effect of socio-economic and biological variables on birth weight in Madhya Pradesh. *Malays J Nutr* 10:159–71

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Daniel Biftu Bekalo¹  · **Zelalem Mehari Nigussie²** · **Belay Birlie Yimer³** · **Zelalem Alamrew Anteneh²**

Zelalem Mehari Nigussie
zelalem2011@gmail.com

Belay Birlie Yimer
belayabirlie@gmail.com

Zelalem Alamrew Anteneh
kzolam@gmail.com

¹ Department of Statistics, Haramaya University, Dire Dawa, Ethiopia

² Department of Epidemiology and Biostatistics, College of Medicine and Health Sciences, Bahir Dar University, Bahir Dar, Ethiopia

³ Department of Statistics, Jimma University, Jimma, Ethiopia