Asian Journal of Advances in Research

4(1): 924-935, 2021

INFANT MORTALITY AND ITS RISK FACTORS IN ETHIOPIA

MIESO GERMAMO WAKO[1,2*], DERBACHEW ASFAW[2] AND ABATE WALDETENSAI[3*]

[1]Bule Hora University, Bule Hora, Ethiopia.
[2]Arba Minch University, Arba Minch, Ethiopia.
[3]Ethiopian Public Health Institute, Addis Ababa, Ethiopia.

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Received: 22 June 2021
Accepted: 28 August 2021
Published: 01 September 2021

Original Research Article

ABSTRACT

Introduction: Information on infant mortality is an important indicator of a country's socioeconomic development; quality of life, help to estimate infant's risk level and support the development of strategies to reduce this risk such as promoting birth spacing.

Objective: The main objective of this study was to identify and explain the effects of the demographic and socio economic determinant factors of infant mortality in Ethiopia

Methods: 2016 Ethiopian Demographic & Health Survey data was used. The data was analyzed with multilevel logistic regression model and maximum likelihood estimation for parameters.

Results: We noticed the data has nested structure; there exists regional disparity in infant mortality. The result of this study revealed that out of the 31,037 infants considered in the analysis, 45.5/1000 was died, while the remaining was alive. The multilevel logistic regression model result showed that region, place of residence, mother education level, source of drinking water, wealth index of household, mothers exposure to media, birth order, breast feeding, age of mothers at first birth and birth interval were found to be the significant risk factors of infant mortality.

Conclusion: From the multilevel logistic regression model, all the three modelsmay be significant, indicating that there is real multilevel variation among infant death in Ethiopia. From the three multilevel model compared, random intercept model gives better result than the null and random slope model upon analyzing data that have nested structure or hierarchical in nature. The intra correlation coefficient is also suggests that there is clear variation of infant death across the region of Ethiopia. The deviance based chi square value is significant for multilevel random intercept model implies that in comparison to the model with multilevel random intercept and fixed slope model the multilevel random intercept model has a better fit the data.

Keywords: Infant death; region of residence; hierarchical models; risk factors.

**Corresponding author: Email: abyw.res31@gmail.com, mieso333germamo@gmail.com;*

1. INTRODUCTION

Infant mortality is an indicator of community health advance that depends on society life expectancy at birth relating with survival chance [1]. Ministry of health of federal democratic republic of Ethiopia claimed a target for infant mortality, to reduce infant mortality rate from level of 44/1000 to 20/1000 by 2019/20. In nearly all populations, deaths before age one comprise the majority of deaths before age five. Attaining Universal health coverage is the direction of Ethiopia's health sectors to ensure standard of health and health system development for Childs with an end to all preventable child deaths by 2035 [2,3]. One in every 17 Ethiopian children dies before the first birthday [4]. In Ethiopia, results from the 2011 EDHS data show a remarkable decline in all levels of childhood mortality. Mortality trends can also be examined by comparing data from DHS surveys conducted in 2000, 2005, and 2011. Infant mortality rates obtained by these surveys evidence a continuous declining trend in mortality [3]. Mortality rates in urban areas are consistently lower than in rural areas, although the difference is quite small for neonatal mortality. Infant mortality is 29 percent higher in rural areas (76 deaths per 1,000 live births) than in urban areas (59 deaths per 1,000 live births) [4].

These rates were highest in Benishangul Gumuz, part of Ethiopia and lowest in Addis Ababa, the capital city of Ethiopia. According to 2016 the EDHS, the neonatal, infant, and under-five mortality rates were 29, 48, and 67 deaths per 1,000 live births respectively [3]. The ratio of infant death was 1 in 35 within the first month and 1 in 21 before celebrating the first birthday with 1 in 15 children death before reaching the fifth birthday [5]. This was also still high when compared to the target of the minimum Millennium Development Goals (MDGs) by 2030 and so studying the factors affecting infant mortality is interesting.

Numerous interventions and efforts regarding primary health care, health education, and extension of health services by government, private sector, and non-governmental organizations intended to improve the lives of children in Ethiopia. However, the lives and health situations of these children, especially the infants have remained poor. The less than 5 mortality rate is higher compared to the target of the minimum Millennium Development Goals (MDGs) by 2030 and regions show large variations in infant mortality. This is also true for the infant mortality that there is large variation of infant death in different region of Ethiopia.

Information on infant mortality is relevant to a demographic assessment of the population and it is an important indicator of a country's socioeconomic development and quality of life. It can also help to estimate how many infants may be at higher risk of death and support the development of strategies to reduce this risk such as promoting birth spacing. For a long time the scholars have been interested in the study of infant mortality to investigate its impact on population change [6,7]. Therefore this study tries to address the regional variation of infant mortality and explore the major risk factors of infant death in Ethiopia.

2. MATERIALS AND METHODS

2.1 Data

The data used for this study was obtained from 2016 Ethiopian demographic Health Survey (EDHS), which was conducted by the Central Statistical Agency (CSA) from January 18, 2016 to June 27, 2016, based on a nationally representative sample that provides estimates at the national and regional levels and for urban and rural areas. According to this survey women aged 15-49 and men aged 15-59 were randomly selected across Ethiopia with a nationally representative sample of 16,650 households and 15,683 female respondents. The sampling frame used for the survey came from the Ethiopian Population and Housing Census (PHC), which was implemented by the Ethiopian CSA in 2007. According to 2007 PHC, the census frame was a complete list of 84,915 enumeration areas (EAs) and enumeration area was a geographic area covering on the average 181 households [4,5]. Infant death before reaching one year was considered as response variable. We dichotomize the variable as 1 if there was died infants before one year and 0 if alive.The explanatory variables were classified into demographic, socioeconomic and health and environments. A demographic characteristic, socio- economic (wealth index, religion, Health status, access to source of information, habit and education level) of all study participants (mothers/women) at all level (individuals and regional) as a factors of infant mortality.

Multilevel analysis methodology was used for analysis of data with complex pattern of variability, with focus on nested source of variability by taking account the variability associated with each level of nesting. Multilevel models allow the relationship between the explanatory variables at different level and dependent variables at lower level to be estimated, enable the extent of variation in the outcome of interest to be measured at each level assumed in the model both before and after the inclusion of the explanatory variables in the model [8]. Two levels of data were stated (for instance individual infants and region) in a multilevel logistic regression model. Units at one level were nested within units at the next higher level (i.e. individual infants were nested in the region). In this study the basic data structure of the two-level logistic regression is a collection of N groups (regions) and within-group j (1, 2... N) a random sample n_j of level-one units (individual infant). The response variable was denoted by Y_{ij} and given by: $Y_{ij} =$ {1 if the i^{th} infant in the j^{th} region was died, 0 if the i^{th} infant in the j^{th} region was alive} with probabilities $\prod_{ij} = P(Y_{ij} = 1/X_{ij}, U_j)$ is the probability of being died for the i^{th} infant in the j^{th} region and $1- \prod_{ij} = P(Y_{ij} = 0/X_{ij}, U_j)$ is the probability of being alive for the i^{th} in the j^{th} regions. Here, Y_{ij} follows a Bernoulli distribution. Like the logistic regression the $\prod_{ij}$ is modeled using the link function, logit. The two-level logistic regression model can be written as,

$$log\,[\textstyle\prod_{ij} / 1- \prod_{ij}] = \beta_o + U_{oj} \quad (1)$$

Where: U_{oj} is the random effect at level two and β_o is the vector of unknown coefficients of the covariates and intercept. Without U_{oj}, Equation (1) can be considered as standard logistic regression model. Therefore, conditional on U_{oj}, the Y_{ij} can be assumed to be independently distributed. Here, U_{oj} is a random quantity and follows $N\,(0,\sigma^2_u)$ [9].

Test of Heterogeneity Proportion: To apply multilevel analysis the first logical step is to test heterogeneity of proportions between groups. Here we present two commonly used test statistics that are used to check for heterogeneity [10]. The two tests used for this purpose are a chi-square based nonparametric test and a parametric test. Now, we present the nonparametric test. To test whether there are indeed systematic differences between the groups, the well known Chi-Square test for contingency table can be used. In this case the Chi-Square test statistic is as follow.

$$X^2 = \sum_{j=1}^{N} n_j \frac{(\hat{\pi}j - \hat{\pi})^2}{\hat{\pi}(1-\hat{\pi})} \sim X^2(N-1 \quad (2)$$

This statistic follows approximately central chi-square distribution with N-1 degrees of freedom [8,11].

Estimations of Between and Within Group Variance: Further note that $\hat{\pi}_j$ is an estimate for the true variance between the group dependent probabilities and an estimator for the variance of $\prod_j$ can be obtained by using [8]:

$$\tau^2 = S^2{}_{between} - \frac{S^2{}_{within}}{\tilde{n}}$$

$$Where: \tilde{n}\ \ is\ \ defined\ \ as: \tilde{n} = \frac{1}{N-1}\left\{M - \frac{\sum_{j=1}^{N} n^2 j}{M}\right\}, \quad (3)$$

$$S^2{}_{between} = \frac{\hat{\pi}(1-\hat{\pi})}{\tilde{n}(N-1)} X^2\ \ and\ \ S^2 within = \frac{1}{M-N}\sum_{j=1}^{N} n_j(1-\pi_j)$$

Where S^2 Population variance, M individual mean, t^2 independent sample (t test), n sample size and N population size.

The empty model for a binary outcome variable refers to a population of groups (level-two units) and specifies the probability distribution for group-dependent probabilities π_j (probability of having i^{th} infant in j^{th} group (region) died before one year of age) and also this model does not consider explanatory variables into account. We focus on the model that specifies the transformed probabilities $f\,(\pi_j)$ to have a normal distribution[12]. This is expressed for a general link function $f\,(\pi)$ and given by the formula:

$$f\,(\textstyle\prod_j) = \beta_o + U_{oj} \quad (4)$$

Where $f\ (\prod_j)$ - is the population average of the transformed probabilities β_o *and* U_{oj} is the random deviation from this average for group *j*. If $f\ (\pi)$ is the logit function, then $f\ (\prod_j)$ is just the log-odds for region *j*. Thus, for the logit link function, the log odds have a normal distribution in the population of groups, which is obtained by:

$$\text{Logit}\ (\textstyle\prod_j) = \beta_o + U_{oj} \tag{5}$$

For the deviations U_{oj} it is assumed that they are independent random variables with a normal distribution with mean zero and variance δ_o^2 .This model decomposes the total variance in the outcome in to two parts i.e. an individual variance captured by the individual level error term, and a group variance, the group level error term. The unconditional model is therefore useful for investigating the amount of variation that exists within versus between groups. One way to quantify this is to calculate the intra-class correlation coefficient (ICC), which represents the proportion of the total variance that is attributable to between group differences and it provides an assessment of whether or not significant between groups variations exists [12]. Then the intra class correlations (ICC) at regions level is given by:

$$p = \sigma^2 u/\sigma^2 u + \sigma^2 e \tag{6}$$

Where;σ^2_u is the between group variance which can be estimated by U_{oj} and $\sigma^2_e = \pi^2/3 = 3.29$

was within regions variance. This model does not include a separate parameter for the individual level variance [13]. This is because the individual level residual variance of the y_{ij} (death or alive of infants) follows Bernoulli distribution directly from the probability of having infant death (π_j) which was given by:

$$Var\ (\mathcal{E}_{ij}) = \pi_j(1 - \pi_j) \tag{7}$$

Let π_o represent the probability corresponding to the average value β_o, then it is defined by:

$$\pi_{o\ =}\ \text{Logit}\ (\beta_o) = \exp\ (\beta_o)/1{=}\exp\ (\beta_o) \tag{8}$$

Or

$$\pi_{ij} = \frac{\exp\left(\beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj}\right)}{1 + \exp\left(\beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj}\right)} \tag{13}$$

As [14] the introduced there was an approximate formula which valid when the variances are small and the approximate relation of between the populations variance is given by:

$$Var(\ \pi_j) \sim \sigma^2/(\ f(\pi_o)^2 \tag{9}$$

For the logit function, this yields:

$$Var(\ \pi_j) = \pi_o(1 - \pi_o))^2\ \sigma_o^2 \tag{10}$$

Random Intercept Model: Here the random intercept model consider only random effect of each indicators of infant mortality it mean that the region differ with respect to the average value of infant death. However, there was no different relation between indicators of infant mortality among groups (regional level). Let consider $X_1, X_2 \ldots\ldots\ldots X_k$ as a predictor's data matrix denoted by X_h (h = 1. 2,……k) these variables are denoted by with their values indicated by X_{hij} [10]. Some or all of those variables could be level one variables, the success probability is not necessarily the same for all individual in a given region.The random intercept model expresses the log-odds, that means the logit of π_{ij} and it was the sum of a linear function of all indicators of infant mortality and also its formula is given as:

$$Logit\left(\pi_{ij}\right) = \beta_{oj} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \ldots + \beta_k X_{kij} + U_{oj} = \beta_{oj} + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj} \tag{11}$$

In this equation the logit(π_{ij}) does not include a level-one residual because it is an equation for the probability of having infant death (π_{ij}) rather than for the outcome y_{ij} . Since, β_{oj} was assumed to vary randomly and it was the sum of an average intercept β_o and region dependent deviations U_{oj} then the logit was given by:

$$Logit\left(\pi_{ij}\right) = \beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj} \tag{12}$$

Where: β_h - is a unit difference between the values of two individuals in the same group is associated with a difference of β_h in their log-odds, or equivalently, a ratio of exp(β_h) in their odds, U_{oj} - is random part of the model and it is assumed that they are mutually independent and normally distributed with mean zero and variance σ_o^2. Random Coefficient Model: The hierarchical data structure can be modeled by using multilevel modeling in order to correct the estimated standard errors and to allow for clustering of observations within units [9]. A random effects model gives the method for estimating the degree of correlation in the outcome that exists at the region level and it represent factors influencing the outcome variable that cannot be quantified in a large-scale social survey. The intercepts β_{oj} as well as the regression coefficients, or slopes, β_{1j} are group (region) dependent. These group dependent coefficients can be split into an average coefficient and the group dependent deviation:

$$\begin{aligned}\beta_{oj} &= \beta_o + U_{oj}\\ \beta_{1j} &= \beta_1 + U_{1j}\end{aligned} \tag{14}$$

Then, by substituting in the above equation the logit (π_{ij}) was given by:

$$\begin{aligned}&Logit(\pi_{ij}) = (\beta_o + U_{oj}) +\\ &(\beta_1 + U_{1j}) X_{1ij} = \beta_o +\\ &\beta 1 X_{1ij} + U_{oj} + U_{1j} X_{1ij}\end{aligned} \tag{15}$$

At group level we have two random effects which were called the random intercept denoted by U_{oj} and the random slope denoted by U_{1j}. Both of them are assumed to have mean zero and the variances are denoted by δ_o^2 and δ_1^2 respectively. Where, β_o - is the average intercept of the response variable, β_1 - is fixed regression coefficient given explanatory variable X_1, U_{oj} - is the random coefficient in the model, $U_{oj} + U_1 X_{1ij}$- is the random part of the model can be considered as interaction by region and predictors (X).The two random effects that characterized region U_{oj} and U_{1j} were correlated and the pairs of random (U_{oj} , U_{1j}) effects were independent and identically distributed. Thus, the variances and covariance of the level-two random effects were (U_{oj} , U_{1j}) denoted by:

$$\begin{aligned}\operatorname{var}(U_{oj}) &= \sigma_{oo} = \sigma_0^2\\ \operatorname{var}(U_{1j}) &= \sigma_{11} = \sigma_1^2\\ \operatorname{cov}(U_{oj}, U_{1j}) &= \sigma_{01}^2\end{aligned} \tag{16}$$

The above single predictor model can be extended by adding more explanatory variable that has random effects on outcome variables. Suppose that there are k level-one explanatory variables X_1, X_2 , ..., X_k, and consider the model where all predictor variables have varying slopes and random intercept. Let

$$\begin{aligned}&\beta_{oj} = \beta_o + U_{oj}\\ &\beta_{1j} = \beta_1 + U_{1j}, \ldots, \beta_{hj} = \beta_k +\\ &U_{hj},\ for\ h = 1, 2\ldots\ldots k,\end{aligned} \tag{17}$$

Then, we have:

$$\begin{aligned}&Logit(\pi_{ij}) = (\beta_o + U_{oj}) +\\ &(\beta_1 + U_{1j}) X_{1ij} + \ldots + (\beta_k + U_{hj}) X_{hij}\\ &= \beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_o + \sum_{h=1}^{k} U_{hj} X_{hij}\end{aligned} \tag{18}$$

Where: $\beta_o + \sum_{h=1}^{k} \beta_h X_{hij}$ -is the fixed part of the model, $U_o + \sum_{h=1}^{k} U_{hj} X_{hij}$ -is the random part of the model, U_{oj} , U_{1j} ….. Uhj - are assumed to be independent between groups but may be correlated within a regions.

3. RESULTS AND DISCUSSION

The results from Binary logistic regression analysis (Table 1) revealed that region, place of residence, mother level of education, source of drinking water, mother religion, mother exposure to mass media, wealth index of household, age of mother at first birth, breast feeding, smoking, birth order, Sex of infants and birth interval were significant predictors of infant death at 5% level of significance. From Table 1 below, the wald value for region was 279.089 and p-value 0.000 which is less than 0.05, this indicates that region was significantly associated with the infant death. The odds of infants death in Amhara region, Oromia region, Benishangul-Gumuz region, SNNPR region, Gambela region, Harari region,Addis Ababa and Dire Dawa were significantly different from Tigray region.

Infants in Ahmara region were 1.282 (OR=1.282) times more likely to die than the residents of Tigray region. Looking carefully into the Death probability in

Oromia (OR=1.416), Gambela (OR=1.572) and Harari (OR=1.302), the probability of infant death was higher than Tigray region. While being in SNNPR (OR=0.89), Benishangul-Gumuz (OR=0.865), Addis Ababa (OR=0.828) and Dire Dawa (OR=0.654) gives a significant advantage compared with the Tigray region. The odds of infant death for mothers who had primary and above education were significantly differing from those who had no education The odds of infant death for mothers who had primary level of education was 2.743 times more likely to alive than that of infant death of mother who had no education. While for mother who had secondary level of education was 2.004 more likely to be alive than mother who do not have no education. The odds of infant death for mother who had higher level of education was 2.283 more likely to be alive than mother who do not have no education. This implies that infants with mothers who had no education have the highest chance of being dying compared to infants with mothers who have primary and above level of education.

The hierarchical structure of the data was formed such that 31,037 individual infants were nested in the 11 geographical regions based on 2016 EDHS data. In empty model we first fit a simple model with no predictors i.e. an intercept-only model that predicts the probability of infant death. The simplest non-trivial specification of the hierarchical linear model is a model in which only the intercept varies between level two units and no predictor (explanatory) variables are entered in the model. The empty model can be considered as a parametric version of assessing heterogeneity among the regions. The estimates of parameters and standard errors are presented in Table 2. The overall mean of infant death was estimated at β_o = -0.2806 and significant. In addition the between region (regional variance) was estimated as $\hat{\sigma}_u^2$ =0.1662. Here the null hypothesis tested is δ_o^2 = 0. i.e., there is no regional variation in infant death in Ethiopia. Based on the results Wald χ^2 = 5.1692 which means $(z-value)^2 = (0.1662/0.0731)^2$, df = 1 with the corresponding p-value= 0.0229, the null hypothesis has to be rejected, indicating strong evidence that the between region variance is non zero. The variance of the random factor is significant which indicates that there are regional differences in infant death and thus, multilevel analysis can be considered as an appropriate approach for further analysis. The variances $\sigma_e^2 = 3.29$ and $\hat{\sigma}_u^2$ =0.1662 estimate the variation among infants or individual and among regions of the country respectively. In variance component model it is possible to decompose variance in to regional level (higher level) and infant level. Infant (level-1) variance was to assess how much of the variation was due to individual themselves and region (level-2) variance was to assess how much of the variation was due to regional level. According to [15] the individual (level-1) variance was fixed to $\frac{\pi^2}{3}$ =3.29 for logit model. In order to get an idea of how much of variation in infant death was attributable to the region level factors, it is useful to see the intra-region correlation coefficient (ICC)= $\frac{0.1662}{0.1662+3.29}$ = 0.048,which measures the proportion of variance of the infant death that is between regions, not within regions. The intra-region correlation coefficient (ICC) in intercept only model is 0.048 which is significant at 5% level of significance. This means that around 4.80% of the variance in infant death is due to variation across (between) regions. Whereas the remaining 95.2% attributable to individual level, within region differences by considering all predictor values zero.

3.1 Random Intercept Model

The random intercept and fixed slope logistic regression model is a multilevel model which have random intercept and fixed coefficients of predictors. To identify the effect of explanatory variables a multilevel binary logistic regression model with random intercept and fixed explanatory variables was estimated using MLwiN software and stata software. The deviance based chi-square test for significance overall goodness fit model (X^2 = -2log(likelihood of Null model)-(-2log (likelihood of final model))=42271.918-39592.892=2679.026 , *df*= 24, P=0.000 indicates that the random intercept model with the fixed explanatory variables was found to be a better fit as compared to the empty model discussed in section above. The results from the random intercept model (Table 3), showed that the random intercept (β_{oj}) is significant implying that the average infant death is differ from region to region. The intercept estimation is random at the regional level $\mathrm{var}(u_{oj})$. Thus, the value of $\mathrm{var}(u_{oj}) = \delta_u^2$ =0.0567 is the estimated variance component of the intercept. The multilevel logistic regression analysis result displayed in Table 3 confirmed the significance of regional difference in infant death. The deviance based chi-square=226.18, *df* = 1, *p*-value=0.000 for random effects in random intercept model, suggesting that infants with the same characteristics in different regions have different death probability.

Table 1. Result of the final Logit regression model on the determinant of infant mortality

Variables	**Category**	$\hat{\beta}$	**S.E.**	**Wald**	***df***	***p*-value**	**Exp ($\hat{\beta}$)**	**95% C.I**	
								Lower	**Upper**
Region	Tigray **(ref)**			279.08	10	.000*			
	Afar	-.009	.064	.021	1	.885	.991	.873	1.124
	Amhara	.249	.064	15.098	1	.000*	1.282	1.131	1.454
	Oromia	.348	.061	32.065	1	.000*	1.416	1.255	1.597
	Somalia	-.032	.056	.319	1	.572	.969	.868	1.081
	Benishangul	-.146	.057	6.475	1	.011*	.865	.773	.967
	SNNPR	-.117	.057	4.197	1	.040*	.890	.795	.995
	Gambela	.452	.062	52.829	1	.000*	1.572	1.392	1.776
	Harari	.264	.074	12.837	1	.000*	1.302	1.127	1.504
	Addis Ababa	-.188	.072	6.799	1	.009*	.828	.719	.954
	Dire Dawa	-.424	.103	16.888	1	.000*	.654	.534	.801
Place residence	Urban (ref)								.
	Rural	-.211	.051	17.275	1	.000*	.810	.733	.894
..Mothers level of education	No **(ref)**			110.47	3	.000*			
	Primary	1.009	.157	41.281	1	.000*	2.743	2.016	3.732
	Secondary	.695	.156	19.793	1	.000*	2.004	1.475	2.722
	Higher	.826	.165	25.084	1	.000*	2.283	1.653	3.154
Drinking water	Piped **(ref)**			10.898	2	.004*			
	Improved	-.170	.059	8.317	1	.238	.843	.751	.947
	Unimproved	.039	.033	1.395	1	.004*	1.040	.975	1.109
Religion	Orthodox **(ref)**			93.476	5	.000*			
	Catholic	.432	.167	6.732	1	.009*	1.541	1.111	2.135
	Protestant	.304	.231	1.733	1	.188	1.355	.862	2.130
	Muslim	.126	.164	.586	1	.444	1.134	.822	1.564
	Traditional	.561	.166	11.437	1	.001*	1.752	1.266	2.425
	Others	.327	.207	2.490	1	.115	1.387	.924	2.082
Mother exposure to media	No **(ref)**								
	Yes	.184	.077	5.767	1	.016*	1.202	1.034	1.398
Wealth index	Poorest **(ref)**			23.744	4	.000*			
	Poorer	.178	.044	16.381	1	.000*	1.195	1.096	1.302
	Middle	.177	.045	15.738	1	.000*	1.194	1.094	1.303
	Richer	.092	.045	4.168	1	.041*	1.097	1.004	1.198
	Richest	.175	.045	15.076	1	.000*	1.191	1.090	1.301
Age	<20 **(ref)**			82.973	2	.000*			
	20-29	-.016	.158	.010	1	.921	.984	.723	1.341
	30-39	-.266	.159	2.825	1	.093	.766	.562	1.045
Breast feeding	No **(ref)**								
	Yes	.547	.025	474.50	1	.000*	1.728	1.645	1.815
Smoke	No **(ref)**								
	Yes	-.435	.106	16.904	1	.000*	.647	.526	.797
Birth order	2-3 **(ref)**			1006.7	2	.000*			
	4-6	-1.169	.038	942.53	1	.000*	.311	.288	.335
	7 and above	-.704	.038	338.74	1	.000*	.495	.459	.533
Sex	Female **(ref)**								
	Male	.074	.024	9.539	1	.002*	1.077	1.028	1.129
Birth interval	One year **(ref)**								
Constant	Two years	0.823	.053	242.48	1	.000*	2.277	2.053	2.525
		0.985	0.29	11.021	1	.001*	0.373		

Table 2. Estimates for empty model

Fixed part	**Estimates**	**Std error**	***z*-value**	***p*-value**	**95%*CI***	
					Lower	**Upper**
β_o =intercept	-0.2806.	0.1236	-2.27	0.023	-0.522	-0.038
Random effect	Estimates	Std. error	*z*-value	*p*-value	Lower	Upper
$\hat{\sigma}_u^2$	0.1662	0.0731	2.2736	0.0229	0.0702	0.3933
ICC(Rho)	0.0481	0.0201				

The results displayed in Table 3 showed that the intra-region correlation coefficient (ICC) is estimated as $\hat{\rho} = \frac{0.0567}{0.0567 + 3.29} = 0.0169$, which is statistically significant at the 5% level of significance. This indicates that about 1.69% of the total variability in infant death is accounted for the differences in regions, with the remaining unexplained 98.31% attributable to individual differences. In addition, type of place of residence, mothers educational level, source of drinking water, religion, mothers exposure to mass media, breast feeding, smoke cigarettes, birth order, sex of infant and birth interval were also found to be significant determinant factor for the variation in infant death.

From the random part the level-one and level-two variances of the random intercept model $\sigma^2{}_\varepsilon$ and $\sigma^2{}_u$ was found to be significant, which implies that individual infants and regions difference contributing for the variation of infant death from the random intercept and fixed explanatory model.The fact that the difference in deviance between the random intercept model and the variance component model is highly significant it means that the random intercept model is a highly significant improvement as compared to the variance component model or there was enough evidence to consider random intercept and fixed explanatory variables model as the best fit as compared to random intercept only model. This is because -2log (likelihood) or deviance by introduction of explanatory variables in the null model was decreased.

3.2 Random Coefficient Model

It is possible to generalize the model so that the effect of level-1 covariates is different in each region. This can be done by adding random coefficients in front of some of the individual-level covariates of the model. In random intercept model we allowed the intercept only to vary across regions by fixing explanatory covariates, but the relation between explanatory and dependent variables can differ between groups (regions in our case). This model contains a random slope for place of residence and mother level of education, which means that it allows the effect of the coefficient of this variable to vary from region to region. This model is more appropriate than the previous model for the variables being used since it was intuitive to assume that infant breast feeding varies from region to region. From the analysis we investigated the value of ρ the intra-region correlation coefficient, in each model. By adding level-1 predictors, the ICC increased and estimated as $\hat{\rho} = \frac{(0.0040 + 0.0533)}{(0.0040 + 0.0533 + 3.29)} = 0.01712$, meaning that roughly 1.712% of the total variability in infant death is attributable to the random factor and region in random coefficient multilevel binary logistic model. From the above Table again the random coefficient estimates for intercepts and the slopes vary significantly at 5% significance level or the confidence interval does not include zero, which implies that there is a considerable variation in the effects of breast feeding of infants and this variables differ significantly across the regions.

The deviance-based Chi-square value of 224.14, p = 0.000, is the difference between the model with and without random effects models. The significance of this difference further indicates that a model with a random coefficient is more appropriate to explain regional variation than a model with fixed coefficients. The correlation between the intercept and random slope of breast feeding was 0.0064.This implies that the death of infants who were breast feeding was less than those who were never breast feeding by a larger factor at regions with higher intercepts compared to regions with lower intercepts (Table 4).

3.3 Model Comparison

The choice of relevant multilevel model is an important step, and it should be based on the necessity of parsimony in the model. This means that models should be as simple as possible [8,9]. The deviance-based chi-square value (χ^2 =500.65 p-value<0.000) for the empty model is shown in ''Table 4''. The deviance-based chi-square is calculated as the difference of log likelihoods between an empty model for single level logistic regression and empty model for multilevel logistic regression, which is to be compared with the critical value from the chi-square distribution with 1 degree of freedom (Table 4). The significance of this test further implies that an empty model with random intercept is more appropriate than an empty model.

The deviance-based chi-square value (X^2 *=226.18, p*-value <0.000) is significant for multilevel random intercept model and the deviance-based chi-square value (X^2=224.14, *p*-value<0.000) for multilevel random slope model (random coefficient model) is also statistically significant. Both models seem to better fit to the data compared to the empty model. However based on the, Deviance, AIC and BIC values, random intercept model was the smallest among the model considered. Therefore the random intercept model best fits the data.

The current study was aimed to model and identify determinant factor for infant mortality using the 2016 EDHS data. Hierarchical models such as: null model, random intercept model, and random coefficient model were used for analyzing infant mortality data. The result shows that region of residence, place of residence, religious belief, age of mother at first birth, mothers education level, source of drinking water, mothers exposure to mass media, breast feeding, smoke cigarettes, birth order, sex of infant and birth interval were the risk factors for infant death. In concordance with a study by [10]),we found a significant effect of the age of mother at first birth on infant death. The likelihood of infant death for age of mothers at first birth in an interval between 20 and 29 was 39.4% lower than the group of mothers aged <20 years. This was even to 35.4% for mothers aged between 30 and 39.

Table 3. Result of random intercept model

Fixed effect Variable	**Category**	$\hat{\beta}$	**S.E**	**z-value**	**p-value**	**Exp(** $\hat{\beta}$ **)**	**95% CI**	
							Lower	**Upper**
Residence **(**Urban)	Rural	0.217	0.050	4.28	0.000*	1.241	0.117	0.316
Mother education	Primary	-0.314	0.034	-9.09	0.000*	0.731	-0.381	-0.246
Level **(**No education)	Secondary	-0.186	0.082	-2.25	0.024*	0.830	-0.348	-0.024
	Higher	-1.014	0.156	-6.46	0.000*	0.363	-1.321	-0.706
Drinking water	Improved	0.2176	0.063	3.44	0.001*	1.242	0.0934	0.341
(Piped)	Unimproved	0.1784	0.059	3.02	0.003*	1.194	0.0626	0.294
Religion (Orthodox)	Catholic	-0.118	0.168	-0.70	0.4881	0.888	-0.449	0.212
	Protestant	-0.296	0.047	-6.25	0.000*	0.744	-0.389	-0.203
	Muslim	0.126	0.037	3.35	0.001*	1.133	0.052	0.199
	Traditional	-0.103	0.133	-0.78	0.438	0.902	-0.364	0.158
	Others	-0.422	0.166	-2.54	0.003*	0.654	-0.748	-0.096
Media (No)	Yes	-0.187	0.076	-2.45	0.014*	0.829	-0.337	-0.037
	Poorer	-0.009	0.037	-0.02	0.980	0.991	-0.074	0.072
Wealth index	Middle	-0.086	0.040	-2.14	0.032*	0.917	-0.165	-0.007
(Poorest)	Richer	-0.004	0.041	-0.11	0.910	0.995	-0.086	0.076
	Richest	-0.179	0.043	-4.11	0.000*	0.835	-0.265	-0.094
Mother age **(**<20)	20-29	-0.025	0.027	-9.11	0.000*	0.975	-0.305	-0.197
	30-39	0.006	0.015	0.04	0.968	1.006	-0.302	0.314
Breast feeding	Yes	-0.546	0.025	-21.78	0.000*	0.579	-0.596	-0.497
Smoke **(**No)	Yes	0.4345	0.105	4.11	0.000*	1.542	0.2273	0.641
Birth order **(**2-3)	4-6	0.4649	0.026	17.85	0.000*	1.589	0.4138	0.516
	>7	1.1683	0.038	30.71	0.000*	3.205	1.0942	1.243
Sex (Female)	Male	-0.074	0.024	-3.09	0.002*	0.928	-0.121	-0.027
Birth interval (1yr.)	Two years	-0.822	0.052	-15.57	0.000*	0.440	-0.925	-0.718
Constant		0.2906	0.105	2.76	0.006*	1.336	0.0840	0.497
Random effect								
$\sigma_u^2 = \mathrm{var}(u_{oj})$		0.0567	0.025				0.0231	0.139
ICC(Rho)		0.0169						

Table 4. Result of random coefficient model

Fixed effect Variable	Category	$\hat{\beta}$	S.E	z-value	p-value	Exp($\hat{\beta}$)	95% CI	
							Lower	Upper
Residence (Urban)	Rural	0.2168	0.050	4.77	0.000*	1.241	0.117	0.316
Mother education	Primary	-0.314	0.034	-9.11	0.000*	0.730	-0.382	-0.247
Level (No education)	Secondary	-0.187	0.082	-2.27	0.023*	0.829	-0.350	-0.025
	Higher	-1.012	0.157	-8.45	0.000*	0.364	-1.320	-0.705
Drinking water (Piped)	Improved	0.218	0.063	3.45	0.001*	1.243	0.094	0.342
	Unimproved	0.179	0.059	3.03	0.002*	1.195	0.063	0.294
Religion	Catholic	-0.112	0.168	-0.67	0.505	0.893	-0.443	0.218
(Ort0hodox)	Protestant	-0.294	0.047	-6.18	0.000*	0.745	-0.387	-0.200
	Muslim	0.133	0.038	3.51	0.000*	1.142	0.059	0.208
	Traditional	-0.098	0.133	-0.74	0.462	0.906	-0.360	0.163
	Others	-0.419	0.166	-2.52	0.012*	0.658	-0.745	-0.092
Media(No)	Yes	-0.188	0.076	-2.45	0.014*	0.828	-0.338	-0.037
Wealth index (Poorest)	Poorer	-0.001	0.037	-0.05	0.963	0.998	-0.075	0.071
	Middle	-0.086	0.040	-2.13	0.033*	0.917	-0.165	-0.006
	Richer	-0.003	0.041	-0.09	0.926	0.996	-0.085	0.077
	Richest	-0.178	0.043	-4.09	0.000*	0.836	-0.264	-0.092
Mother age (<20)	20-29	-0.252	0.027	-9.16	0.000*	0.777	-0.307	-0.198
	30-39	0.007	0.157	0.05	0.962	1.007	-0.301	0.316
Breast feeding (No)	Yes	-0.555	0.033	-16.60	0.000*	0.574	-0.621	-0.490
Smoke (No)	Yes	0.432	0.105	4.09	0.000*	0.862	0.225	0.639
Birth order (2-3)	4-6	0.464	0.026	17.85	0.000*	1.589	0.413	0.516
	>7	1.168	0.038	30.69	0.000*	3.205	1.094	1.243
Sex (Female)	Male	-0.074	0.024	-3.09	0.002*	0.928	-0.121	-0.027
Birth interval (1yr)	Two years	-0.822	0.052	-15.56	0.000*	0.440	-0.925	-0.718
Constant		0.2874	0.103	2.77	0.006*	1.331	0.0837	0.491
Random effect								
$\sigma_u^2 = \text{var}(u_{oj})$	0.0040	0.0055					0.0002	0.0576
$\text{Var}(u_{1j})$	0.0533	0.0242					0.0214	0.1324
$\text{Cov}(u_{1j},u_{oj})$	0.0064	0.0084					-0.010	0.0230
ICC(Rho)	0.0171							

Table 5. Model comparison with other models

Fitted model	Null model	Random intercept Only model	Random coefficient model
-log likelihood	-21135.959	-19796.446	-19797.447
Deviance-based chi-squar	500.65	226.18	224.14
p-value	0.000	0.000	0.000
Model fit diagnostics			
Deviance	42271.918	39592.892	39594.894
AIC	42275.92	39644.89	39650.93
BIC	42292.62	39861.81	39884.54

Infant death was also associated with mother education level. The odds of death for infants who live with mother who have primary education were 2.743 times more likely to alive when it is compared with those mother who do not have education and the log odds of infant alive also increased by 2.004 for mother with secondary and increased by 2.283 for mother with higher education when compared with mother who do not have school education at all. Generally, infant death is likely to decrease when mother educational level increases. These result also agree with the results of several studies as in [8,10] Place of residence of the infant is significant determinants of infant death. Infants who live in rural areas were more likely to be die than those who live in urban areas. This result is similar with the results found by [16,17]. The study also indicated that wealth index of household negatively related with the death of infants, indicating that infants who were from poorest household were more likely to die than those who were from richest household. This result coincides with findings by [17].

The random intercept model is significantly different from zero indicating that infant death differs from region to region. The deviance based chi-square test for random effects in random intercept model is also high (X^2 =500.65, *df* =1, *p*-value<0.001). This indicate that the random intercept model with the fixed slope is found to give a good fit as compared to the empty model for predicting infant death across regions of Ethiopia.The variance component of random intercept is also large which further supports the fact that there is a high variability in infant death across regions. Within region variation further implies that death of infants within regions are more likely (or heterogeneous) than between regions. Thus multilevel analysis has demonstrated that different regions have significantly different mean effects, and that the effect for place of residence is different in rural and urban areas across the regions. This was supported in the finding by [8]. On the other hand it was found that the Random intercept and explanatory variables provide additional information. First, the variances of the random components related to the random term were found to be statistically significant implying the presence of differences in death of infant across the regions. Secondly, from explanatory variables considered here, the effect of breast feeding varies from region to region. Third, the interaction between random parts of breast feeding provide significant differences on infant death across regions A study done by [10] on under-five child mortality found similar results.

4. CONCLUSION

Infant mortality in Ethiopia is one of the challenging problems that the country needs to address. Place of residence, mother educational level, age of mother at first birth, source of drinking water, wealth index of house hold, religious beliefs, breast feeding, sex of infant, birth order and birth interval were the determinants for death of infant. The deaths of infants are also significantly affected by smoking, religious beliefs (Muslims, protestant, orthodox and others) and media following place of residence, mother's education level, source of drinking water, living standard and breast feeding. Sex and age difference contributing for variation of infant deaths. Generally, from the three multilevel model compared, random intercept model gives better result than the null and random slope model upon analyzing data that have nested structure or hierarchical in nature. That is the multilevel random intercept include the random parts that shows the variation of individual level and group level factors**.** The death of infants from mother who have education at least primary, secondary and higher school attainments are less than that mother who do not have school education at all. This may be because of the fact that educated mother are well informed about family planning and how to give care for their infants. The effect of regional variations for breast feeding, further implies that there exist considerable deference in infant death among regions and a model with a random intercept was more appropriate to explain the regional variation than a model with fixed coefficients or without random effects. There should be an increased health care centers and community awareness about family planning and it is useful to give attention that a lot of effort needs to be made in family planning programs to give awareness on early marriage, spacing birth interval to reduce infant death. The study also revealed that there was high infant's death that use unimproved source of drinking water. Therefore, government should take a measure of action on preparing improved source of drinking water to tackle this problem. Generally, government should give more support and emphasis on those regions with high rates of infant death. In order to decreases death levels in regions with lower levels, the socio-economic status of the regions has to be raised. As a consequence, differences in the level of infant death between regions would be reduced.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. UNICEF. Committing to Child Survival : A Promise Renewed Progress Report 2013. New York, NY 10017, USA; 2013.
2. UNICEF. Humanitarian Situation Report. 2018.
3. MoH M of H. National Strategy for Newborn and Child Survival in Ethiopia National Strategy for Newborn and Child Survival in Ethiopia. Addis Ababa, Ethiopia; 2019.
4. CSA, EDHS. 2016 Demographic and Health Survey Key Findings [Internet]. Addis Ababa; 2016. www.DHSprogram.com
5. Bekele S, Getachew Y. Survival Analysis of Under-Five Mortality of Children and its Associated Risk Factors in Ethiopia. J Biosens Bioelectron. 2016;7(3).
6. Special Programme for Research and Training in Tropical Diseases (TDR). Steategic review

of traps and targets for Tsetse and African Trypanosomiasis control. Geneva, Switzerland; 2004.
Report No.: TDR/IDE/TRY/05.1.

7. Mathews TJ, Driscoll AK, Ph D. Trends in Infant Mortality in the United States , 2005 – 2014. 2017;(279):1–8.
8. Bedane AS. Multilevel Logistic Regression Applications. 2017;(March 2016).
9. Goldstein H. Multilevel Statistical Models [Internet]. London; 1999.
Available:http//:www.arnoldpublishers.com/support/goldstein.htm
10. Senbeta BA. Multilevel Logistic Modeling Under Child Mortality variation among regional state of Ethiopia. Hawassa; 2012.
11. Ferede T. Multilevel Modelling of Modern Contraceptive Use among Rural and Urban Population of Ethiopia. Am J Math Stat. 2013;3(1):1–16.
12. Miao G. Application of Hierarchical Model in Non-Life Insurance Actuarial Science. Mod Econ. 2018;9:393–9.
13. Tom, A. B Roel JB. Multilevel Analysis_ An Introduction to Basic and Advanced Multilevel Modeling-. New Delhi: SAGE; 1999.
14. Berger MPF, Berkhof J, Browne WJ, Busing FMTA, Leeuw J de, Draper D. Handbook of Multilevel Analysis [Internet]. Jande, Leeuw Erik M, editor. New York: Springer Science+ Business Media, LLC; 2008. springer.com
15. Snijders TAB. Multi level Analysis. Statistics D of, editor. UK: University of Oxford; 2012.
16. Baraki AG, Akalu TY, Wolde HF, Lakew AM. Factors affecting infant mortality in the general population : evidence from the 2016 Ethiopian demographic and health survey (EDHS); a multilevel analysis. BMC Pregnancy Childbirth. 2020;20(299):1–8.
17. Wheatley L, Wheatley L, Middleton C. Factors affecting child mortality. University of Tennessee at; 2015.
