Statistical Analysis of Neonatal Mortality: A Case Study of Ethiopia

Berhanu Teshome Woldeamanuel*
Department of Statistics, Mekelle University, Ethiopia

Abstract

Background: Neonatal mortality is a significant public health concern worldwide. It is estimated that four million neonatal deaths occur annually, 98% of which occur in developing countries. In Ethiopia neonatal mortality is a major problem accounting for more than 42% of under-five deaths. This study is an attempt to study the determinants of neonatal mortality in Ethiopia using data collected in Ethiopian demographic and health survey.

Methods: The survey collected information from a total of 16,515 women aged 15-49 years out of which 9209 women were considered in this study. To meet our objectives, descriptive statistics and Poisson, negative binomial, zero inflated Poisson and zero-inflated negative binomial regression models were used for data analysis considering household, maternal and socio-demographic, socio-economic and environmental variables as explanatory variables and number of neonatal deaths per-woman as the response variable.

Results: Each of the four models was compared by a variety of statistical techniques and it was found that zero-inflated negative binomial model was a better fit than the other models. Based on descriptive statistics results 23.2% of mothers have experienced at least one neonatal death in their lifetime. From result of the zero-inflated negative binomial regression model, being born to mothers whose age at first birth is at least 20 years, whose level of education is secondary and above, who reside in urban areas and who attended at least four antenatal care visits significantly decreases the risk of neonatal mortality.

Conclusion: Neonatal mortality must decline more rapidly to achieve the Millennium Development Goal (MDG) target for under-five mortality in Ethiopia. Increasing access to maternal and child health services in rural areas, improving the level of education of mothers, encouraging utilization of antenatal care services and improving access to safe/pipe drinking water are recommended as possible interventions.

Keywords: Neonatal mortality; Poisson; Negative binomial; Zero-inflated Poisson; Zero-inflated; Negative binomial

Abbreviations: AIDS: Acquired Immune Deficiency Syndrome; AIC: Akaike Information Criteria; ANC: Antenatal Care; BIC: Bayesian Information Criteria; CI: Confidence Interval; CSA: Central Statistical Agency; DHS: Demographic and Health Survey; GLM: Generalized Linear Model; LRT: Likelihood Ratio Test; MDG: Millennium Development Goal; ML: Maximum Likelihood; NB: Negative Binomial; NM: Neonatal Mortality; NMR: Neonatal Mortality Rate; OR: Odds Ratio; PR: Poisson Regression; USAID: United States Agency for International Development; WHO: World Health Organization; ZINB: Zero-Inflated Negative Binomial; ZIP: Zero-Inflated Poisson

Introduction

The neonatal period (birth to 28th day of life) is the most vulnerable and high-risk time in life because of the highest mortality and morbidity incidence in human life during this period. Neonatal mortality refers to the probability of dying of a live born infant within the first 28 days of life. An estimated 40 percent of deaths in children less than five years of age occur during the first 28 days of life [1]. The remaining 60 percent of deaths occur during the subsequent 1800 days of the first five years of life. The average daily mortality rate during the neonatal period is close to 30 fold higher than during the postnatal period (one month to one year of age). During 2010, an estimated 7.7 million children under five years of age died worldwide. This included 3.1 million neonatal deaths, 2.3 million post neonatal deaths (age one month to one year) and 2.3 million childhood deaths (age 1-5 years) [2].

Despite the very high burden of mortality, the problem of neonatal mortality has received little attention until relatively recently. There is now a growing consensus within the international community that increased efforts are needed to reduce newborn deaths if further progress is to be made in reducing child mortality

Maternal health before, during and after pregnancy, conditions at the time of labour and delivery and post-natal care of babies play a significant role in reducing neonatal mortality. The major direct causes of neonatal deaths vary from region to region and from country to country. Countries with high estimates of Neonatal Mortality Rate (NMR) (NMR>45/1,000 live births) have a larger proportion of deaths (nearly 50%) that is attributable to infections such as pneumonia, diarrhea and tetanus.

In Ethiopia mortality trends can be examined by comparing data from DHS conducted in 2000, 2005 and 2011. Infant and under-five mortality rates obtained by these surveys evidence a continuous declining trend in mortality. On the other hand, even though neonatal mortality rate decreased from 49 deaths per 1,000 live births in 2000 to 39 deaths per 1,000 live births in 2005, it has remained stable at 37 deaths per 1,000 live births in 2011. Approximately 42% of the under-5 mortality in Ethiopia is attributable to neonatal deaths. According to the 2011 Ethiopia Demographic and Health Surveys (DHS), the country is experiencing a high neonatal mortality rate at 37 per 1000 live births, comparable to the average rate of 35.9 per 1000 live births for the African region overall [3,4].

*Corresponding author: Berhanu Teshome Woldeamanuel, Department of Statistics, Mekelle University, Ethiopia. Tel: +251910118464; E-mail: beretshome19@gmail.com

Received: February 09, 2018; Accepted: April 23, 2018; Published: April 30, 2018

Citation: Woldeamanuel BT (2018) Statistical Analysis of Neonatal Mortality: A Case Study of Ethiopia. J Preg Child Health 5: 373. doi:10.4172/2376-127X.1000373

Copyright: © 2018 Woldeamanuel BT. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
Neonatal Mortality (NM) has become a significant public health concern worldwide. Of the 130 million babies born every year, a total of 4 million die in the neonatal period. In Africa, the neonatal mortality rate is high (41 neonatal deaths per 1000 live births) and the distribution in the Sub-Saharan region (Eastern, Western and Central Africa) range between 42 and 49 neonatal deaths per 1000 live births. Every year about 300,000 African babies die on the first day of their birth. Like many other countries in the Sub-Saharan regions, babies born in Ethiopia are at risk of dying in the neonatal period, it has remained stable at 37 deaths per 1,000 live births in 2011 [5].

In Ethiopia, there are limited studies that look at the importance of the association between each determinant and NM. In addition, to our knowledge no studies have used the recent Demographic Health Survey that was conducted in Ethiopia in the year 2011 regarding this problem. Moreover, NM in Ethiopia is an important concern, where it is essential for monitoring the current health programs and formulating policies for improving the current situation. From that concrete recommendations may be given to policy makers and health program managers for improving the health policies and the maternal and child care strategy. An understanding of the factors related to neonatal mortality is therefore important to guide the development of focused and evidence-based health interventions to prevent neonatal deaths. Therefore the purpose of this study is to identify and analyse the various possible risk factors of neonatal mortality in Ethiopia using count data models.

### Materials and Methods

#### Data

The current study is based on Ethiopia Demographic and Health Survey 2011 data which is part of the worldwide measure DHS project funded by the United States Agency for International Development (USAID) conducted by Central Statistical Agency (CSA). The survey was designed to provide estimates for the health and demographic characteristics in the country. Nationally 16,515 women of age 15–49 were included in the survey.

This study is based on retrospective data from women of childbearing aged (15–49). Information on neonatal mortality is obtained from mothers asking relevant questions about her live births, when the child was born and whether the birth she gave was still alive. In case the baby is dead, the number of day when was he or she died was recorded within the first month of birth, where neonatal mortality refers to children who died within the first 28 days of their birth. Finally, response on 9,209 women of age 15–49 who gave birth were analysed based on count regression model to determine the predictors of neonatal mortality in the region.

#### Variables in the study

The dependent variable of the study is number of neonatal death that a mother has experienced counted as 0, 1, 2,... The most important expected determinants factors of neonatal mortality from their theoretical justification included in this study are given in Table 1 below (Table 1).

| Variable Name | Variable description | Category |
|---------------|----------------------|----------|
| AGEMOTH | The age of the mother at the time of the first birth (in years) | 0=less than 20  1=20 and above |
| EDUMOTH | The highest educational level the woman attained | 1=no education  2=primary  3=secondary and higher |
| EDUHUSB | Husband/partner's educational attainment | 1=no education  2=primary  3=secondary and higher |
| ANC | Number of antenatal visits during pregnancy | 1=no visits  2=1-3 visits  3=4 visits and above |
| PNC | Postnatal care after delivery | 0=no  1=yes |
| RESIDENCE | Place of residence | 0=urban  1=rural |
| RELIGION | Religion | 1=Coptic orthodox  2=protestant  3=Muslim |
| REGION | Geographical region/area | 1=Tigray  2=Afar  3=Amhara  4=Oromiya  5=Somali  6=Benchishangul-Gumuz  7=SNNP  8=Gambela  9=Harari  10=Addis Ababa  11=Dire Dawa |
| PDELIV | Place of delivery | 0=home  1=health facility |
| SWATER | Source of water supply | 1=pipe  2=protected source  3=unprotected source |
| TOILET | Availability of toilet facility | 0=no facility  1=with facility |
| OCUMOTH | Woman’s occupation | 1=Not working  2=agricultural sector  3=non agriculture sector |

Table 1: Description of covariates.
Methodology

Modeling count variables is a common task in public health research and social sciences. The classical Poisson regression model for count data is often of limited use in these disciplines because empirical count data sets typically exhibit over-dispersion and/or an excess number of zeros. Hence in modeling count data, it is a usual practice after the development of Poisson regression model to proceed with analysis of correcting for overdispersion if it exists. One of the approaches to modeling overdispersion is to use Quasi Likelihood estimation technique [1,6]. Besides the problem of overdispersion, it is also common that real life count data exhibit excess zeros. While the negative binomial model can deal with overdispersion, the negative binomial regression models can be used to handle excess zeros in their own way. In almost all practical cases, the count data are skewed, non-negative, overdispersed and have excess zeros. These features have motivated the application of various methods and models for count data regressions. Hence, this paper provides a practical approach of modeling count data focusing on data that exhibits overdispersion and excess zeros [6,7].

Count data models

One of the crucial questions in statistical analysis of count data is how to formulate an adequate probability model to describe observed variation of counts. Because of the restrictive nature of equidispersion assumption in standard Poisson model, researchers have developed techniques and tests that allow detecting the overdispersion (or under-dispersion) in the population. Recently, Zero-inflated models have been developed to take into account the excess of zeros in the data. The Zero-inflated model can be seen as a finite mixture model where one distribution is considered as a degenerate process with a unit point mass at zero [8].

Poisson regression model (PR)

Consider we have n independent random variables denoted by Y1, Y2, ..., Yn. Where Yi denote neonatal death for ith mother in her life time. Poisson regression model for count data assumes that each observed count y is drawn from a Poisson distribution with rate of occurrence of an event, parameter λ, for i = 1, 2, ..., n. Let Xs denote a vector of covariates in the study for the ith mother.

Then the Poisson equation of the model with rate parameter λ is given by:

\[ p(Y_i = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} , \lambda_i > 0 \text{ and } y_i = 0,1,2,... \]  

(1)

The mean of the response variable λ is related with the linear predictor through the so called link function [9].

Now let X be nx (k+1) matrix of explanatory variables. The relationship between Yi and kth row vector of X, Xs, linked by g(λs), is the canonical link function given by:

\[ E(Y) = \lambda = e^{Xs\beta} \]  

(2)

where, Xs=(x1, x2, ..., xk) is the ith row of covariate matrix, (with x=1 indicating a vector of unity that represent the values for intercept term) and β= (β0, β1, ..., βk) is unknown (k+1) dimensional vector of regression parameters.

The above two equations are called Poisson regression model or log-linear model.

Negative binomial regression model (NBR)

The main problem in the use of standard Poisson regression model, where there a more dispersed observations or if the data have excess zeros, another way of modeling such over-dispersed count data is a negative binomial (NB) distribution which can arise as a gamma mixture of Poisson distributions.

Let Y1, Y2, ..., Yn be a set of n independent random variables one parameterization of its probability density function is, then the negative binomial model given as:

\[ p(Y_i = y_i) = \frac{\Gamma(y_i + 1)}{\Gamma(y_i + k) \Gamma(k)} \left( \frac{\alpha}{\alpha + \lambda_i} \right)^y_i \left( \frac{\lambda_i}{\alpha + \lambda_i} \right)^k , y_i = 0,1,2,... \text{ and } \alpha \geq 0 \]  

(3)

Covariates can be introduced into a regression model based on the NB distribution via the relationship:

\[ \ln(\lambda_i) = \beta + \sum_{j=1}^{k} x_{ij} \beta_j \]  

(4)

where xij is the ith observation corresponding to the jth covariate, k is the number of covariates in the model and \( \beta_j \) is the jth regression parameter [10,11].

Zero-inflated regression models

One characteristic of the Poisson distribution is that the mean of the distribution is equal to the variance; however when there are excess zeros, probability of zero in the standard model will be less than the expected. The problem of standard models in underpredicting zeros and overpredicting ones is very common and sometimes this problem can be very severe when there are a lot of zeros in the distribution. In such cases, Zero inflated Poisson (ZIP) and Zero inflated negative binomial (ZINB) models can be used to account for excess zeros. The zero values in the ZIP model can be viewed as comprising two parts. One portion of the zero counts arises from the inflated part of the distribution and the other portion comes from what would be expected given a Poisson distribution with parameter λ [12].

Zero-inflated Poisson regression model (ZIP): Generally, the zero-inflated probability mass function has the form:

\[ p(Y = y) = \begin{cases} \phi + (1-\phi) \cdot p(Y = 0) & \text{if } y = 0 \\ (1-\phi) \cdot p(Y = y) & \text{if } y = 1,2,... \end{cases} \]  

(5)

If Yi are independent random variables having a zero-inflated Poisson distribution, the zeros are assumed to arise in two ways corresponding to distinct underlying states. The first state occurs with probability Φ, and produces only zeros, while the other state occurs with probability (1-Φ) and leads to a standard Poisson count with mean λ. In general, the zeros from the first state are called structural zeros and those from the Poisson distribution are called sampling zeros. This two-state process gives a simple two-component mixture distribution with probability mass function:

\[ p(Y = y) = \begin{cases} \phi + (1-\phi) \cdot e^{-\lambda_i} \frac{\lambda_i^y}{y!} & \text{if } y = 0 \\ (1-\phi) \cdot e^{-\lambda_i} \frac{\lambda_i^y}{y!} & \text{if } y = 1,2,... \end{cases} \]  

(6)

This is denoted by Y~ZIP (λ, Φ) such that 0 ≤ Φ<1, where λ is the mean of the non-zero outcomes that can be modelled with the associated explanatory covariates using a natural logarithmic link function as:

\[ \ln(\lambda_i) = Xs\beta \]  

(7)

where Xs=(1, x1, x2, ..., xk) is a (k+1) x1 vector of explanatory variable of the ith subject and β is (k+1)x1 vector of regression coefficient.
parameters. $\Phi_i$ ($0 < \Phi_i < 1$) is the probability of an excess zero (being in the zero mortality state) determined by a logit model. To predict membership in the "Always Zero" group, we can use the same variables or we can use a smaller subset of the variables or even different variables altogether, that is:

\[
\ln\left(\frac{\varphi_i}{1 - \varphi_i}\right) = \gamma \gamma + \text{equivalently: } \varphi_i = \frac{\exp(\gamma \gamma)}{1 + \exp(\gamma \gamma)} \quad i=1,2,\ldots,n \quad (8)
\]

The deviance (log likelihood ratio statistic) is given by Long [13]:

The overall model goodness of fit may be symbolically stated as:

\[
\sum_{i=1}^{n} y_i \log \left(\frac{\hat{y}_i}{\hat{y}_i} + \hat{\lambda}_i \right) = 0 \quad (9)
\]

where $\lambda_i$ is the mean of the non-zero response that can be modelled with the associated explanatory covariates using a natural logarithm link function as defined in equation (7) and $\Phi_i$ is the probability of excess zeros which can be estimated by the logistic regression as defined in equation (8) [12].

The ZINB model is a special case of a two-class finite mixture model like the ZIP model with mean $E(Y) = \lambda_i (1 - \Phi_i)$ and variance $\text{Var}(Y) = \lambda_i (1 - \Phi_i)$ ($1 + \alpha \Phi_i + \Phi_i$), where the parameters $\lambda_i$ and $\Phi_i$ depend on the covariates and $\alpha \geq 0$ is a scalar. Thus we have over-dispersion whenever either $\Phi_i$ or $\alpha$ is greater than 0. Thus, the equation in (10) reduces to NB when $\Phi_i = 0$ and to the ZIP when $\alpha = 0$.

The likelihood function for ZINB model is:

\[
L = \Pi_{y_i=0} \left[\varphi_i + (1 - \varphi_i) \frac{1}{1 + \alpha \lambda_i} \right] \Pi_{y_i>0} \left[\frac{\Gamma(y_i + 1)}{\Gamma(y_i)} \left(\frac{\lambda_i}{1 + \alpha \lambda_i}\right)^{y_i} \right] \quad (10)
\]

\[
\hat{L} = \hat{\Phi}_i \hat{\lambda}_i \hat{\alpha}_\lambda \quad (11)
\]

Assessing model fit

Overall regression test: Often after fitting a model it is good to assess how model is good fit. The null and alternative hypothesis for overall model goodness of fit may be symbolically stated as:

\[
H_0: \hat{\beta}_1 = \hat{\beta}_2 = \hat{\beta}_3 = \cdots = \hat{\beta}_k = 0 \text{ versus } H_a: \text{Not all } \hat{\beta}_i = 0 \text{ for } j=1, 2, 3, \ldots, k
\]

This can be tested using the deviance test with $k$ degrees of freedom. The deviance (log likelihood ratio statistic) is given by Long [13]:

\[
D = 2 \left\{ \sum y_i \log \left(\frac{\hat{y}_i}{\hat{y}_i} + \hat{\lambda}_i \right) \right\} \quad (12)
\]

We reject the null hypothesis of no model goodness of fit for large value of the test for a given level of significance.

Test for significance of a single variable: The null hypothesis for this test may be stated as "Factor $X_i$ does not have any value added to the prediction of the response given that other factors are already included in the model."

\[
H_0: \hat{\beta}_i = 0
\]

To test such a null hypothesis, one can use

\[
t_i = \frac{\hat{\beta}_i}{\text{se}(\hat{\beta}_i)} \quad (13)
\]

where $\hat{\beta}_i$ is the estimated regression coefficient and $\text{se}(\hat{\beta}_i)$ is the estimate of the standard error of $\hat{\beta}_i$. This test statistic has the t-distribution with $n-k-1$ degree of freedom.

Test for over dispersion parameter: The negative binomial regression model reduces to the Poisson regression model when the over dispersion parameter is not significantly different from zero. To assess the adequacy of the negative binomial model over the Poisson regression model, we can test the hypothesis:

\[
H_0: \alpha = 0 \text{ versus } H_a: \alpha > 0.
\]

This is a test of significance of the over dispersion parameter $\alpha$. The presence of the over dispersion parameter $\alpha$ in the NB regression model is justified when the null hypothesis $H_0: \alpha = 0$ is rejected. In order to test the above hypothesis a score test statistic is proposed. The general score test statistic for testing $H_0$ can be given by:

\[
S_n = \frac{\left\{ \left( y_i - \hat{\lambda}_i \right)^2 - y_i \right\}^2}{2 \hat{\lambda}_i \hat{\lambda}_i} \quad (14)
\]

where $\hat{\lambda}_i$ is the predicted value from the Poisson regression model. Under the null hypothesis that the data follow a Poisson model, the limiting distribution of the score statistic is chi-squared with one degree of freedom [14,15].

Goodness of fit tests: In this section, several goodness-of-fit measures will be briefly discussed, including the, likelihood ratio test, Vuong statistic, Akaike Information Criteria (AIC) and Bayesian Information Criteria (BSC).

Likelihood ratio: The advantage of using the maximum likelihood method is that the likelihood ratio test may be employed to assess the adequacy of the negative binomial over the Poisson because negative binomial will reduce to the Poisson when the dispersion parameter, $\alpha$, equals zero. In this study a likelihood ratio was used to compare the Poisson with the negative binomial and zero-inflated Poisson with zero-inflated negative binomial since Poisson is nested on negative binomial and zero-inflated Poisson is nested in zero-inflated negative binomial. However this will not be used to compare Poisson or negative binomial with the zero inflated Poisson and negative binomial as long as these models are not nested one on the other.

The likelihood ratio statistic is given by:

\[
T = 2 \left\{ \ell_1 - \ell_0 \right\} \quad (15)
\]

where $\ell_1$ and $\ell_0$ are the model’s log likelihood under the alternative and null hypothesis, respectively. $T$ has a chi-square distribution with one degrees of freedom. This method is not appropriate for models which are not nested. In such situations, we will use another method such as the Akaike information criteria (AIC) and Bayesian information criteria (BIC).

Akaike information criteria (AIC)

AIC is the most common means of identifying the model which fits
well by comparing two or more than two models. It tries to balance the goodness of fit against the complexity of the model. It is given by:

$$AIC = -2\ell + 2k$$  

where $$\ell$$ denotes the log likelihood of a model that is to be compared with the other models and $$k$$ is the number of parameter in the model including the intercept. A good model is the one which has the minimum AIC value [16,17].

The BIC is defined as:

$$BIC = -2\ell + k \log(n)$$  

where $$\ell$$ denotes the log likelihood of the model, $$n$$ is the sample size and $$k$$ is the number of parameters in the model including the intercept. For this measure, the smaller the BIC, the better the model is [18].

Vuong statistic

Vuong has introduced a test which is a well suited method to compare zero-inflated regression models to standard non nested model for counts data. Suppose $$f_i(y/x)$$ and $$f^*_i(y/x)$$ denote the probability density functions of zero-inflated model (ZIP or ZINB) and parent (traditional) model (PO or NB), respectively. We want to test the following hypotheses:

$$H_0:$$ Two distribution functions are equivalent

$$H_a:$$ Two distribution functions are different

Now we define

$$m_i = \log \left( \frac{\hat{f}^*_i(y/x)}{\hat{f}_i(y/x)} \right),$$  

(17)

Where $$\hat{f}^*_i(y/x)$$ and $$\hat{m}_i = \sum_{i=1}^{n} m_i$$ are predicted probabilities of the corresponding models, respectively. Let $$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m_i$$ and $$s_m = \left[ \frac{1}{n}\sum_{i=1}^{n} (m_i - \overline{m})^2 \right]^{1/2}$$ denote the mean and standard deviation of the measurements $$m_i$$. Then the Vuong statistic is defined as:

$$V = \frac{\overline{m}}{s_m} \frac{\log(n)}{\sqrt{n}}$$  

(18)

For large sample size and under the null hypothesis, the statistic $$V$$ has an asymptotic standard normal distribution. Within the family of ZIP models, testing if a Poisson model is adequate corresponds to testing:

$$H_0: \Phi=0 \text{ vs. } H_a: \Phi>0$$

Where possible tests include the likelihood ratio test (LRT) and the score test. For a general ZIP regression model, the LRT for zero-inflation is given by:

$$R_e = -2\left[ l(\hat{\lambda}_0) - l(\hat{\lambda}_0, \hat{\phi}) \right]$$  

(19)

Where $$l(\hat{\lambda}_0)$$ and $$l(\hat{\lambda}_0, \hat{\phi})$$ are, respectively, the maximized log-likelihoods under the Poisson regression and the ZIP regression models [16,19].

Results

Table 2 below shows various descriptive statistics of neonatal deaths. The result shows that 76.8% of the mothers have not faced any neonatal death in their lifetime. If we observe the overall pattern of neonatal death at the national level, it is highly skewed to the right with excess zeroes (skewness=3.353, kurtosis=17.419) (Table 2).

Table 3 presents summary statistics of the variables that are assumed to affect neonatal mortality. The variables included are region, place of residence, mothers’ education level, husband/partner’s education level, number of antenatal visits, place of delivery, postnatal care, mothers’ occupation, religion, mothers age at first birth, source of water supply and availability of toilet facility (Table 3).

The total number of women considered in this study was 9209 of which 2140 of them experienced neonatal death. Of the total number of neonatal deaths per woman, 14.3% and 25.9% of neonatal deaths have occurred in urban and rural areas, respectively. Another maternal variable that possibly has a strong bearing on the survival prospects of a child is the mother’s age at the time of first birth. Regarding mothers’ age at first birth, 5827 (63.3%) of children were from mothers of age less than 20 years while the remaining 3382 (36.7%) were from mothers with age 20 years and above. Of these, 26.6% and 14.8% of children have died before the age of one month, respectively. Of the total of 9209 women included, 5987 (65%) have delivered at home and only 3038 (33%) have taken postnatal care. 26.5% of children from mothers who attended postnatal care have died before the age of one month. This figure is 21.6% for children whose mothers have not attended postnatal care.

Overall, more than three fourth 7126 (77.4%) of the respondents (mothers) live in rural areas, while less than one-fourth 2083 (22.6%) of them live in urban areas. From a theoretical perspective, place of residence is an important determinant of child survival. Mothers living in urban areas have a higher chance of getting health service and are aware of the benefit of medication than mothers who reside in rural areas. Table 3 reveals that the percentage of neonatal death for mothers who reside in rural and urban areas are about 26 and 14.3, respectively.

Table 3 also reveals that there is a decreasing trend in neonatal deaths with regards to education level of mothers (that is, neonatal deaths decrease as education level increases). In particular, the percentage of neonatal mortality was 27.4% for those with no education, 16.6% for those with primary education and 7.1% for those with secondary and higher education. Similarly fathers’ education shows higher mortality for neonates from illiterate fathers (27.8%) and is lower for fathers with secondary and higher education (9.3%).

If we consider the regional variation in neonatal death, Amhara and Benishangul-Gumuz had the highest proportion of neonatal death (29.1% and 28.8%, respectively), while Addis Ababa city administration has the lowest proportion of neonatal death (only 8.4%). As far as the numbers of antenatal visits are concerned, the neonatal death percentage decreased
with an increase in the number of visits. Specifically, the neonatal death rate was 26.5% for women who did not attend antenatal care, 19.9% for those who had between one and three visits and 14.7% for those who attended four or more visits (the minimum number recommended by WHO).

Results in Table 3 also show the summary statistics of sanitation indicators - source of water and type of toilet facility. Piped water supply reduces neonatal mortality directly by reducing the incidence of diarrheaa that arises from the ingestion of contaminated water and food and indirectly when caregivers are able to devote more time to childcare instead of water collection activities. The results indicate that 3426 (37.2%) households had piped water while 3912 (42.5%) used water from unprotected source. Children to mothers with unprotected water supply accounted for 19% of neonatal deaths, while those to mothers with piped source of water supply accounted for 19% of neonatal deaths (Table 3).

### Statistical data analysis and model comparison

The variable of interest in this study was the number of neonatal deaths per woman through her life time. Such data can be well fitted by the count data models rather than the linear regression model. Likelihood ratio test (LR), Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used to compare the candidate models to identify the most appropriate model.

### Development of count data regression models for determinants of neonatal mortality: Model identification

An initial step in the model building process is to identify sets of explanatory variables that have the potential for being included in the linear component of a multivariable count data regression model. We start with fitting univariable Poisson regression model. Based on univariable Poisson model we can identify candidate covariates to be considered for the multivariable model. Inclusion of such covariates will be based on a significant reduction in -2LogL, where the bigger the reduction of this value, the better the fit. Including highest educational level of women, husband/partner's educational attainment, number of antenatal visits, postnatal care, place of residence, region, place of delivery, age of mothers at first birth, religion, source of water supply, availability of toilet facility and occupation of mothers in the model, one at a time, result in a significant reduction in -2LogL using the chi-square test. Hence all these covariates are considered in the multivariable count data regression models.

### Table 3: Summary statistics of some important socio-economic, demographic and environmental variables that are assumed to affect the neonatal death in Ethiopia

| Covariates                          | Category                          | Number of neonatal mortality per woman |
|-------------------------------------|------------------------------------|----------------------------------------|
|                                     | 1 (Tigray)                         | 794 (159, 46, 14, 9, 22.3, 1022 (11.1%)) |
|                                     | 2 (Amhara)                         | 687 (122, 41, 19, 6, 21.5, 875 (9.5%))  |
|                                     | 3 (Afar)                           | 673 (245, 84, 16, 14, 29.1, 1232 (13.4%)) |
|                                     | 4 (Oromiya)                        | 959 (235, 62, 20, 10, 25.4, 1286 (14.1%)) |
|                                     | 5 (Somali)                         | 473 (96, 33, 8, 6, 23.2, 616 (6.7%))    |
|                                     | 6 (Benishangul-Gumuz)              | 540 (144, 52, 10, 12, 28.8, 758 (8.2%)) |
|                                     | 7 (SNN)                            | 864 (210, 47, 21, 6, 24.7, 1148 (12.5%)) |
|                                     | 8 (Gambella)                       | 510 (88, 31, 10, 3, 20.6, 642 (7%))     |
|                                     | 9 (Harari)                         | 452 (78, 20, 5, 6, 19.4, 561 (6.1%))    |
|                                     | 10 (Addis Ababa)                   | 466 (38, 4, 1, 1, 8.4, 509 (5.5%))      |
|                                     | 11 (Dire Dawa)                     | 451 (78, 22, 7, 2, 19.5, 560 (6.1%))    |
|                                     | Type of place of residence         | 1786 (234, 48, 10, 5, 14.3, 2083 (22.6%)) |
|                                     | 1 (Urban)                          | 5283 (1258, 394, 121, 70, 25.9, 7126 (74.7%)) |
|                                     | 2 (Rural)                          | 4524 (1152, 376, 114, 64, 27.4, 6230 (67.7%)) |
|                                     | 3 (Secondary and higher)           | 585 (39, 5, 1, 0, 7.1, 630 (6.8%))      |
|                                     | Source of water supply              | 2770 (490, 125, 26, 19, 15.1, 3426 (37.2%)) |
|                                     | 1 (Pipe)                           | 1419 (309, 104, 26, 13, 24.2, 1871 (20.3%)) |
|                                     | 2 (Protected source)               | 2880 (693, 213, 79, 47, 26.4, 3912 (42.5%)) |
|                                     | 3 (Unprotected source)             | 3214 (689, 230, 75, 42, 24.4, 4250 (46.2%)) |
|                                     | Availability of toilet facility     | 3855 (803, 212, 56, 33, 22.3, 4959 (53.8%)) |
|                                     | 1 (No facility)                    | 2702 (553, 161, 36, 27, 22.3, 3479 (37.8%)) |
|                                     | 2 (Protestant)                     | 1355 (277, 76, 31, 12, 22.6, 1751 (19%)) |
|                                     | 3 (Muslim)                         | 3012 (662, 205, 64, 36, 24.3, 3979 (43.2%)) |
|                                     | Age of mothers at 1st birth         | 2477 (1074, 323, 93, 60, 26.6, 5827 (63.3%)) |
|                                     | 20 (and above)                     | 2883 (418, 119, 38, 15, 14.8, 3382 (36.7%)) |
|                                     | Number of antenatal visits during pregnancy | 4375 (1070, 335, 106, 66, 26.5, 5952 (64.3%)) |
|                                     | 2 (1-3 visits)                     | 1316 (240, 67, 15, 4, 19, 164 (17.8%)) |
|                                     | 3 (visits and above)               | 1378 (182, 40, 10, 5, 14.7, 1615 (17.5%)) |
|                                     | Place of delivery                  | 4639 (960, 267, 80, 41, 22.5, 5987 (65%)) |
|                                     | 1 (Home)                           | 2430 (532, 175, 51, 34, 24.6, 3222 (35%)) |
|                                     | 2 (health facility)                | 4637 (956, 259, 78, 41, 21.6, 6171 (67%)) |
|                                     | 3 (postnatal care after delivery)  | 2232 (538, 183, 53, 34, 26.5, 3038 (33%)) |
|                                     | 1 (No)                             | 3453 (886, 294, 91, 59, 27.8, 4783 (51.9%)) |
|                                     | 2 (Yes)                            | 2569 (498, 126, 34, 15, 20.8, 3242 (35.2%)) |
|                                     | Husband/partner's education level   | 1074 (108, 22, 6, 1, 9.3, 1184 (12.9%)) |
|                                     | 1 (No education)                   | 3551 (708, 215, 71, 33, 24.7, 4578 (49.7%)) |
|                                     | 2 (Primary)                        | 1476 (389, 118, 32, 22, 27.5, 2037 (22.1%)) |
|                                     | 3 (Secondary and higher)           | 2042 (395, 109, 28, 20, 21.3, 2594 (28.2%)) |
|                                     | Woman's occupation                 | 7069 (1492, 442, 131, 75, 23.2, 9209 (100%)) |

Table 3: Summary statistics of some important socio-economic, demographic and environmental variables that are assumed to affect the neonatal death in Ethiopia.
Initially, we fitted Poisson model to identify the risk factors of neonatal mortality. The fitted Poisson model is then tested for overdispersion. However, the overdispersion might be due variation among observations or to excess zeroes. This brings the negative binomial model and the zero-inflated models into the picture. Thus, in order to select an appropriate model which fits the data well, four different models were considered, namely: the standard Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial models.

We start here first by checking the overall goodness of fit using deviance (likelihood ratio) statistic. Accordingly, the deviance-based chi-square test provided a chi-square value of 776.51 (p<0.0001) which would imply good fit for the model the fitted Poisson model. The following table display SAS and STATA output of the Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial regression models fit statistics (Table 4).

The score test statistic for over dispersion is $\chi^2=1322.77$ with p-value<0.0001, indicating that there is over-dispersion and the Poisson model is inappropriate. This is also supported by the Pearson Chi-square value divided by the degrees of freedom is close to one compared to the Poisson regression model, which one possible indication that there is over-dispersion.

Next we fit a negative binomial regression model with the same explanatory variables. The likelihood ratio test was used to compare the fit of the standard Poisson with NB regression model. This statistic was found to be statistically significant, indicating that NB regression model is a better fit to the data than the standard Poisson regression model ($\chi^2_{105}=655.11$, p-value<0.0001). Thus, the observed data are better explained by the negative binomial than the Poisson model. Further as one can be seen from Table 4, the NB regression model is a better fit than the Poisson model since it has the smallest AIC (13289.57) as well as BIC (13458.0) values.

Still the reason for over dispersion is unknown thus, we need to fit the zero inflated count models. After fitting zero inflated models, we test $H_0$: $\alpha=0$ versus $H_1$: $\alpha>0$ and $H_2$: $\Phi=0$ versus $H_3$: $\Phi>0$ to identify whether the over dispersion is due to the presence of excess zeros or high variability in the non-zero outcomes. If the null hypothesis testing the inflation parameter $H_2$: $\Phi=0$ is not rejected, then the negative binomial model is appropriate and the over dispersion problem is due to the presence of high variability in the non-zero outcomes. However, if both parameters are significantly different from zero, then the zero inflated negative binomial regression models are more appropriate to fit the data.

Our result indicates that the over dispersion parameter ($\alpha$) in the ZINB regression model is significantly different from zero ($\chi^2=136.61$ with P-value<0.001) and the likelihood ratio chi-square test statistic for testing the inflation parameter is also significant ($\chi^2=600.317$, P-value<0.001).

Hence there was a high variability in the non-zero outcomes and the data excess of zeros. Thus, it is be better to use models which take into account both the high variability due to non-zero outcomes and excess of zeros simultaneously. Figure 1 below plots of difference between the observed probability of each count and the prediction from each of the four models. This result shows that ZINB distribution has the ability to reproduce the zeroes in the population better than ZIP distribution, which means that the ZINB model gives a good prediction of the number of neonatal deaths.

![Figure 1: Plots of difference between the observed probability of each count and the prediction from each of the four models.](image)

| PRM | BIC=13955.15 | AIC=13755.56 | Prefer | Over | Evidence |
|-----|--------------|--------------|--------|------|----------|
| Vs NBRM | BIC=13496.28 | AIC=13289.57 | NBRM | PRM | Very strong |
| Vs ZIP | BIC=13701.6007 | AIC=13302.436 | ZIP | PRM | Very strong |
| Vs ZINB | BIC=13458 | AIC=13266 | ZIP | PRM | Very strong |
| NBRM | BIC=13496.28 | AIC=13289.57 | NBRM | ZIP | Very strong |
| Vs ZIP | BIC=13701.6007 | AIC=13302.436 | ZIP | ZIP | Very strong |
| Vs ZINB | BIC=13458 | AIC=13266 | ZIP | ZIP | Very strong |
| ZIP | BIC=13701.6007 | AIC=13302.436 | ZIP | ZIP | Very strong |
| Vs ZINB | BIC=13458 | AIC=13266 | ZIP | ZIP | Very strong |

Table 4: Tests and fit statistics for model comparisons.
The other measure used to identify the appropriate model in this paper is the Vuong’s test. The Vuong test compares the ZIP model to the standard Poisson model by testing the null hypothesis that both models are equally similar to the observed distribution. The resulting statistic was statistically significant (V=9.17, p-value<0.0001), demonstrating that the ZIP model more reflects the observed data than standard Poisson due to the presence of excess zeroes. Similarly, the Vuong test that was used to compare ZINB versus negative binomial yields a test statistic of V=4.73 with p-value<0.0001. Thus, the Zero-inflated negative binomial regression model more accurately fits the number of neonatal deaths data compared to the standard negative binomial model.

Finally goodness of fit test for the fitted zero-inflated negative binomial regression model was assessed using the Pearson chi-square statistic. Accordingly, the Pearson’s chi-square test provided a chi-square value of 14144.59 (p-value<0.0001) which would imply good fit for the Poisson regression model. Therefore, zero-inflated negative binomial (ZINB) regression model was found to be the most appropriate model which fits the data better than the other possible candidate models. Thus, in the analyses discussed below, the ZINB model is used (Figure 1).

Results in Table 5 provide estimates of the effect of some selected characteristics of mothers on the mortality of their neonate. The first set of coefficients is from the equation predicting counts for “not always zero” group. These coefficients can be interpreted the same way as regular negative binomial coefficients. The factors region, place of residence, education level of mothers, source of water supply, religion, age of mother at time of first birth, antenatal care, postnatal care and education level of husband/partner are statistically significant predictors of this count outcome. Availability of toilet facility, place of delivery and occupation of mothers are not statistically significant (Table 5).

Concerning the regional disparity in neonatal mortality, the results in table 5 show that children in Amhara (OR: 1.65, 95% CI: 1.165–2.337), Benishangul-Gumuz (OR: 1.86, 95% CI: 1.307–2.65), Somali (OR: 1.59, 95% CI: 1.104–2.303), Oromiya (OR: 1.611, 95% CI: 1.142–2.273), Gambella (OR: 1.742, 95% IC: 1.187–2.555), SNNP (OR: 1.573, 95% CI: 1.096–2.257), Harari (OR: 1.54, 95% CI: 1.063–2.227), Tigray (OR: 1.525, 95% CI: 1.068–2.178) regions and Dire Dawa city administration (OR: 1.50, 95% CI: 1.038–2.166) are at a higher risk of experiencing neonatal death than children in Addis Ababa. The study

| Covariates | B       | SE       | Z       | Sig.    | Exp (B)  | 95% C.I for Exp (B) |
|------------|---------|----------|---------|---------|----------|---------------------|
|            | Lower   | Upper    |         |         |          |                     |
| REGION (Addis Ababa) |         |          |         |         |          |                     |
| Tigray | 0.421972 | 0.181799 | 2.32    | 0.020   | 1.5250   | 1.2907 1.6498       |
| Affar   | 0.255199 | 0.185402 | 1.38    | 0.005   | 0.89554  | 0.84569 0.84569     |
| Amhara | 0.500625 | 0.177572 | 2.8227  | 0.007   | 1.6107   | 1.5943 1.6370       |
| Oromiya | 0.476696 | 0.175607 | 2.49    | 0.013   | 1.8612   | 1.7613 1.9614       |
| Somali | 0.466451 | 0.187623 | 2.45    | 0.001   | 1.5732   | 1.4718 1.6857       |
| Benishangul-Gumuz | 0.621206 | 0.180291 | 2.9     | 0.014   | 1.7418   | 1.6345 1.8505       |
| SNN     | 0.453113 | 0.184230 | 2.64    | 0.005   | 1.5388   | 1.4996 1.5840       |
| Gambella | 0.554933 | 0.195842 | 2.29    | 0.022   | 1.3147   | 1.2398 1.3930       |
| Harari | 0.431022 | 0.188525 | 1.14    | 0.150   | 1.1239   | 1.0380 1.2166       |
| Dire Dawa | 0.405160 | 0.187673 | 1.8     | 0.031   | 1.2315   | 1.1229 1.3420       |
| RESIDENCE (Urban) |         |          |         |         |          |                     |
| Rural   | 0.273618 | 0.094689 | 2.89    | 0.004   | 1.3147   | 1.2398 1.3930       |
| EDUMOTH (No educ.) |         |          |         |         |          |                     |
| Primary | -0.38121 | 0.070889 | -5.39   | <0.001  | 0.68303  | 0.36256 0.594665     |
| Secondary and higher | -1.0146 | 0.180223 | -6.63   | <0.001  | 0.36256  | 0.25468 0.516164     |
| SWATER (Pipe) |         |          |         |         |          |                     |
| Protected | 0.116840 | 0.081250 | 1.44    | 0.150   | 1.1239   | 1.0380 1.2166       |
| Unprotected | 0.208672 | 0.072379 | 2.88    | 0.004   | 1.1239   | 1.0380 1.2166       |
| TOILET (No facility) |         |          |         |         |          |                     |
| With facility | -0.08331 | 0.063562 | -1.31   | 0.190   | 0.92007  | 0.81229 1.04213     |
| RELIGION (Orthodox) |         |          |         |         |          |                     |
| Protestant | 0.275364 | 0.098950 | 2.76    | 0.006   | 1.3170   | 1.2257 1.4183       |
| Muslim   | 0.203519 | 0.070923 | 2.61    | 0.009   | 1.3170   | 1.2257 1.4183       |
| AGEMOTH (less than 20) |         |          |         |         |          |                     |
| 20 and above | -0.28676 | 0.077869 | -4.72   | <0.001  | 0.75069  | 0.66635 0.845966     |
| ANC (No ANC visits) |         |          |         |         |          |                     |
| 1-3 visits | -0.22346 | 0.085546 | -2.58   | <0.001  | 0.79975  | 0.76900 0.829745     |
| 4 visits and above | -0.38713 | 0.084462 | -5.62   | <0.001  | 0.79975  | 0.76900 0.829745     |
| PDELIV (Home) |         |          |         |         |          |                     |
| Health facility | -0.11033 | 0.111473 | -0.99   | 0.322   | 0.89554  | 0.71977 1.11422      |
| PNC (No) | 0.366108 | 0.106920 | 3.42    | 0.001   | 1.4421   | 1.1694 1.77832       |
| EDUHUSB (No educ.) |         |          |         |         |          |                     |
| Primary | -0.27480 | 0.082792 | -4.38   | <0.001  | 0.75972  | 0.70738 0.819217     |
| Secondary and higher | -0.26999 | 0.124573 | -2.17   | 0.030   | 0.76338  | 0.58005 0.974493     |
| OCUMOTH (not working) |         |          |         |         |          |                     |
| Agricultural | -0.01470 | 0.072727 | -0.2    | 0.840   | 0.98541  | 0.85449 1.13638      |
| Non-agricultural | -0.001494 | 0.086633 | 1.38 | 0.168   | 1.0963   | 0.96208 1.24925      |
| Intercept | -1.4693 | 0.190754 | -7.70   | <0.001  | 2.89     | 2.89 2.89           |
shows that there is no significant difference in the risk of neonatal death between Afar region and Addis Ababa city administration.

Children born to a mother with secondary or higher schooling was associated with a 64% decreased risk of neonatal death compared to being born to mothers with no education, while primary education of the mothers decreases the risk of neonatal mortality by 32%, keeping all other covariates constant. Results in Table 5 indicate that children born from mothers whose age at first birth is greater than or equal to 20 have a significantly lower risk of mortality compared to those born from mothers whose age at first birth is less than 20 years. The risk of neonatal death was about 25% lower for births to mothers aged 20 and above compared to those aged less than 20. The results also indicate that the chance for membership in the "always zero" group (no neonatal death) increases by a factor of 3.67 for mothers whose religion is Protestant compared to Orthodox. The odds for mothers whose age at first birth is greater than or equal to 20 have a significantly lower risk than mothers with no antenatal care visits. Moreover, mothers with four or above antenatal care visits have a 32% lower risk than mothers with no antenatal care visits.

The second set of coefficients on the bottom half of table 5 predict the dichotomous outcome of group membership (i.e., "not always zero" versus "always zero" groups). The results show that women who delivered at health facility are less likely than those delivered at home to be in the "always zero" group. The odds for health facility delivery to be in the always zero group are 92% lower than for delivery at home. The odds for mothers with access to toilet facility to have no NM (being in always zero group) are 61% lower than those without availability of toilet facility. The results also indicate that the chance for membership in the "always zero" group (no neonatal death) increases by a factor of 3.67 for mothers whose religion is Protestant holding all other variables constant (Table 5).

**Discussion**

The study has empirically investigated and identified the factors that are associated with the risk of neonatal mortality in Ethiopia based on EDHS 2016 data using four count regression methods. The study use retrospective study design. The main limitation of this study is it was difficult to gather all necessary variables and some variables did not get in a way we need. This may affect the association of outcome variable and the determinant factors. The confounding effect of unmeasured variables cannot be controlled. Some determinant factors like mode of delivery, birth weight and maternal smoking habits are a factors associate to neonatal mortality. But even if they are important determinant factors, we fail to include them in the analysis due to incompleteness of the data. Thus, any reader of this manuscript should take in to account the above limitations.

The findings of the study show that being born to a mother with...
education schooling was associated with decreased risk of neonatal death compared to being born to mothers with no education. Thus, educational level of mothers is an important and significant factor of neonatal mortality risks in Ethiopia. In literature most of the studies also report a negative relationship between neonatal death and mothers’ education level. They indicated that education improves the ability of mothers to implement simple health knowledge and facilitates their capacity to manipulate their environment including health care facilities, interact more effectively with health professionals, comply with treatment recommendations and keep their environment clean. Furthermore, educated women have greater control over health choices for their children [23-25].

Results also shows that children born from mothers whose age at first birth is greater than or equal to 20 have a significantly lower risk of mortality compared to those born from mothers whose age at first birth is less than 20. Children living in the rural areas had an increased risk to death compared to those children living in the urban areas. Theoretically, all things being equal, living in urban areas should be associated with a higher standard of living, better sanitation and better health facilities, among other things. Place of residence was found to have an association with neonatal mortality, such that children living in the rural areas had an increased risk to death compared to those living in the urban areas. The risk of perinatal death was less for a child whose mother resides in urban areas compared to their rural counterparts. Theoretically, all things being equal, living in urban areas should be associated with a higher standard of living, better sanitation and better health facilities, among other things. This finding is in agreement with studies by which shows a significant association between place of residence, age at first birth and neonatal death [20-22,26].

Lack of clean water supply combined with little or no health care knowledge can provide routes for infections. The findings of this study show that mothers who use water from unprotected sources are at a higher risk of experiencing neonatal death than those who use pipe water. There is no significant difference in neonatal mortality among those using piped water and protected source of water supply.

Finally, husband education was found to be a significant predictor of neonatal mortality. Further result shows that children of mothers whose husband has primary education and secondary and higher education have a significantly lower neonatal mortality risk than children of mothers whose husband has no education. This finding is consistent with [24,27]. As an indicator of health care service utilization during pregnancy, antenatal care service factors demonstrated a significant association with neonatal mortality.

Conclusion

The study has empirically investigated and identified the factors that are associated with the risk of neonatal mortality in Ethiopia based on Ethiopian DHS 2011 data using methods of count data analysis [5]. Descriptive statistics and multivariable analyses were done to achieve this. In this study, after fitting four different count models it was found that NB, ZINB, regression models were better fitted the data than Poisson and ZIP. Furthermore the zero-inflated negative binomial model was better fitted with data which is characterized by excess zeros and high variability in the non-zero outcome.

According to the results of multivariable ZINB regression model it was found that the variables age of the mother at the time of the first birth, highest educational level the mother attained, number of antenatal visits during pregnancy, husband/partner’s educational attainment, postnatal care-seeking care after birth, place of residence, region, source of water supply and religion are the significant determinants of neonatal mortality in Ethiopia. However, place of delivery, access to toilet facility and mothers’ occupations were found to be insignificant factors of neonatal mortality in Ethiopia.

In recent decades, Ethiopia has achieved significant declines in under-five and infant mortality rates. However, neonatal mortality rates have stayed higher than post-neonatal rates. Further interventions have to be done to reduce mortality rates among neonates. The findings presented in this study have the following research and policy implications:

- The survival experience of children in rural areas is much lower than those in urban areas. Thus, there is a need to increase access to maternal and child health services in rural areas.
- Improving the level of education of mothers is vital. Education of mothers plays an important role in survival of neonate and thus, health programs need to focus on supporting women with little or no education.
- Encourage, advocate and promote utilization of antenatal care services and the importance of following for both the mother and the newborn.
- The study revealed that children of households who use unprotected water as a source of drinking are at a higher risk of neonatal death. Thus, efforts should be made to improve access to safe/pipe drinking water and effective programs to reduce early childbearing of women should be implemented so as to decrease neonatal mortality.

Acknowledgement

We wish to acknowledge Macro International, USA for allowing us access to the 2011 Ethiopian Demographic and Health Survey dataset.

References

1. WHO (2011) Guidelines on optimal feeding of low birth weight infants in low and middle income countries.
2. Rajaratnam JK, Marcus JR, Flaxman AD, Wang H, Levin-Rector A, et al. (2010) Neonatal, post-neonatal, childhood and under-5 mortality for 187 countries, 1970-2010: A systematic analysis of progress towards Millennium Development Goal 4. Lancet 375: 1988-2008.
3. CSA, ORC Macro (2005) EDHS. Addis Ababa, Ethiopia and Calverton, Maryland, USA.
4. Oestergaard MZ, Inoue M, Yoshida S, Mahanani WR, Gore FM, et al. (2011) Neonatal mortality levels for 193 countries in 2009 with trends since 1990: A systematic analysis of progress, projections and priorities. PLoS Med.
5. Morrow AL, Dawoud A (2004) Influencing birth outcomes in Nepal. Lancet 364: 914-915.
6. Karlaftis MG, Tarko AP (1998) Heterogeneity considerations in accident modeling. Accid Anal Prev 30: 425-433.
7. McCullagh P, Nelder JA (1989) Generalized linear models, monographs on statistics and applied probability. 2nd edition, Chapman and Hall, London, UK.
8. Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34: 1-14.
9. Nelder JA, Wedderburn RWM (1972) Generalized linear models. J R Stat Soc 135: 370-384.
10. Lawless JF (1987) Negative binomial and mixed Poisson regression. Can J Stat 15: 209-225.
11. Cameron AC, Trivedi PK (1998) Regression analysis of count data. Cambridge University Press, UK.
12. Natalia S, Gerstel N (2004) Explaining the gender gap in help to parents: The importance of employment. J Marriage Fam 66: 431-451.
13. Long JS (1997) Regression models for categorical and limited dependent variables. Thousand Oaks, CA: Sage.

14. Dean C, Lawless J F (1989) Tests for detecting over dispersion in poisson regression models. J Am Stat Assoc 84: 467-472.

15. Yau Z (2006) Score tests for generalization and zore-inflation in count data modeling. Unpublished Ph.D Dissertation, University of South Caroline, Columbia.

16. Agresti A (2002) Categorical data analysis. Wiley Interscience, New York.

17. Akaike H (1973) Information theory and extension of the maximum likelihood principle. In: Petrov BN, Csaki F (Eds), Second international symposium on information theory, Akademiai Kiado, Budapest, pp: 267-281.

18. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6: 461-464.

19. Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57: 307-333.

20. Oji OS (2008) Risk factors for perinatal mortality in Nigeria: The role of place of delivery and delivery assistant.

21. Hidalgo LA, Chedraui PA, Chevaz MJ (2005) Obstetrical and neonatal outcome in young adolescents of low socio-economic status: A case control study. Arch Gynecol Obstet 271: 207-211.

22. Rutstein SO (2000) Factors associated with trends in infant and child mortality in developing countries during the 1990s. Bull World Health Organ 78: 1256-1270.

23. Chemichovsky D, Meesook OA (1986) Utilization of health services in Indonesia. J Soc Sci Med 23: 611-620.

24. Gabrysch S, Campbell OM (2009) Still too far to walk: Literature review of the determinants of delivery service use. BMC Pregnancy Childbirth 9: 34.

25. Van-Ginneken J, Lob-Levyt J, Gove S (1996) Potential interventions for preventing pneumonia among young children in developing countries: Promoting maternal education. Trop Med Int Health 1: 283-294.

26. Stanton, B, Langsten R (2000) Morbidity and mortality among Egyptian neonates and infants: Rates and associated factors. Ann Trop Med Parasitol 94: 817-829.

27. Mats M (2010) Who can save the unseen? Studies on neonatal mortality in Quang Nihn province, Vietnam. Ph.D. dissertation, Uppsala University.