Bayesian Multilevel-modelling of Acute Respiratory Infections among Under Five Children in Ethiopia

By:

Amanuel Mengistu

A Thesis Submitted to Jimma Univeristy, College of Natural Sciences, Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biostatistics

March 5, 2020

Jimma,Ethiopia

Bayesian Multilevel-modelling of Acute Respiratory Infections among Under Five Children in Ethiopia

By:

Amanuel Mengistu

Main advisor: Prof. Kadi Adiveppa

Co-advisor: Demeke Kifle (Msc)

March 5, 2020

Jimma,Ethiopia

STATEMENT OF THE AUTHOR

I hereby declare that the thesis is my original work and all sources of materials used for the thesis have been duly acknowledged with proper citation. This thesis has not been submitted to any other institution for award of any academic degree, diploma or certificate. I have submitted this thesis to Jimma University in the partial fulfillment for the requirements of Degree of Master of Science in Biostatistics.

The thesis can be deposited in the university library to be made available to borrowers for reference. Brief quotations from this thesis are allowed without requiring special permission if an accurate acknowledgement of the source is made. Requisites for extended quotations for the reproduction of the thesis in whole or in part may be granted by the head of the department of statistics when in her or his judgment the proposed use of the material is for a scholarly interest. In all other instances, however permission must be obtained from the author.

Amanuel Mengistu	....	....
Name	Signature	Date

APPROVAL SHEET

DEPARTMENT OF STATISTICS, SCHOOL OF GRADUATE STUDIES JIMMA UNIVERSITY.

This is to certify that the thesis entitled **"Bayesian Multilevel-modelling of Acute Respiratory Infections among Under Five Children in Ethiopia"** submitted in partial fulfillment of the requirement for the degree of Master of Science in Biostatistics to the college of natural science Jimma University, and is record of original research carried out by **Amanuel Mengistu Merera, ID.No: RM1059/2010**, under my supervision and no part of the thesis has been submitted for another degree or diploma. The assistance and the help received during the course of this investigation have been duly acknowledged. Therefore, I recommended that would be accepted as fulfilling the thesis requirement.

Prof. Kadi Adiveppa	. . . .	. . . .
Name	Signature	Date
Demeke Kifle (MSc)	. . . .	. . . .
Name	Signature	Date

As the members of the board of examiners of MSc. thesis open defense examination of **Amanuel Mengistu Merera**, we certify that we have read and evaluated the thesis and examined the candidate. We recommend that the thesis has been accepted as it fulfills the requirements for the degree of Master of Science in Biostatistics.

. . . .	. . . .	. . . .
Name of External Examiner	Signature	Date
. . . .	. . . .	. . . .
Name of Internal Examiner	Signature	Date
. . . .	. . . .	. . . .
Name of Chairman	Signature	Date

Contents

List of Tables

List of Figures

Acknowledgment

Above all, I would like to thank the almighty God for the gift of health, wisdom and strength throughout any steps for the achievement of my thesis and more of my entire life.

I want to express a sincere acknowledgment to my main advisor Prof. Kadi Adiveppa and co-advisor Mr. Demeke Kifle for their constructive comments and guidance through writing this thesis that contributed to the success of the thesis.

I take this opportunity to sincerely express my gratitude to my beloved family who are sources of pride and encouragement throughout my life. I am thankful to my mother Buze Jiku and my father Mengistu Merera, for their encouragement and prayer to the success of this work.

Finally, I would like to acknowledge Jimma University and all my instructors and staffs of Statistics, Department at Jimma University for their cooperation from starting title selection up to the completion of this thesis and my special appreciation and acknowledgment goes to Mizan-Tepi University for offering me the opportunity and financial support during my study at Jimma University.

Acronyms

ALRI:	Acute Lower Respiratory Infections
ARI:	Acute Respiratory Infection
AIC/BIC:	Akaike/Bayesian Information Criteria
CSA:	Central statistics Agency of Ethiopia
DIC:	Deviance Information Criterion
EA:	Enumeration Areas
EDHS:	Ethiopian Demographic Health Survey
FMoH:	Federal Ministry of Health
ICC:	Intra class Correlation Coefficient
IVAC:	International Vaccine Access Center
LBW:	Low Birth Weight
LRTI:	Lower Respiratory Tract infection
MCMC:	Markov Chain Monte Carlo
MHA:	Metropolis-Hastings Algorithm
MOH:	Ministry of Health
OR:	Odds Ratio
UNICEF:	United Nations Children Fund
URIs:	Upper Respiratory Tract Infections
WHO:	World Health Organization
WLF:	World Lung Foundations

Abstract

Introduction: Acute respiratory tract infection (ARI) is a leading cause of morbidity and mortality in children under the age of 5 years throughout the world. In Ethiopia the prevalence of ARI among under five children was 7% according to 2016 EDHS report.

Objective: This study was aimed to determine factors associated with the prevalence of acute respiratory infection (ARI) among under-five children in Ethiopia using Bayesian multilevel approach.

Methods: This study has been conducted in Ethiopia based on data obtained from EDHS 2016 which mainly focused on under-five children (age from 0-59 months). The survey was collected from a total of 10,641 children aged less than five years out of which 9,918 children were considered in this study. The study was used Bayesian Multilevel Logistic Regression Model to investigate the major risk factors and regional variations in ARI among under five children in Ethiopia. To determine the posterior marginal, the MCMC methods with non-informative priors have been applied. Deviance information criterion was used to compare models.

Results: The analysis showed that the overall ARI prevalence rate between 9,918 under five children was 8.4%. Based on DIC, Bayesian multilevel logistic regression of random coefficient model was found to be more appropriate than empty and intercept model. According to the output of the model childs age, household wealth index, women's educational level, vitamin A supplement, history of diarrhea, maternal work, stunting and source of drinking water were found to be significant factors for prevalence of ARI among children under five years.

Conclusions: This study revealed that there is a significant variation of incidence of ARI among under-five children between and within the regions of Ethiopia. The study recommends all regional states to make remedial measures on public health policy and improve the abilities of stakeholder living in their region toward those major factors associated with the prevalence of ARI among under five children.

Key words: Acute Respiratory Infections, Bayesian Multilevel, MCMC

1 Introduction

1.1 Background of the Study

Acute Respiratory Infections (ARIs) has been defined as one who has a cough, is breathing faster than usual with short, quick breaths or is having difficulty breathing, excluding children that have only a blocked nose (UNICEF, 2016). Acute Respiratory Infections (ARIs) are a group of diseases that includes pneumonia, influenza, and respiratory syncytial virus, and result in 4.25 million deaths worldwide every year (World Lung Foundation, 2011).

Acute respiratory tract infection (ARI) is a leading cause of morbidity and mortality in children under the age of 5 years throughout the world. Pneumonia is one of the most serious manifestations of ARI. Each year, ARI causes 15% of all deaths in children under the age of 5 years globally. About 50% of these deaths occur in Sub-Saharan Africa (UNICEF, 2015).Globally, lower respiratory infections caused more than 2 to 6 million deaths, attributing to the fifth leading cause of death overall and the leading cause of death in children below five years of age (Mortal, 2015). The incidence and prevalence of ARIs are a great burden on low and middle income countries in comparison to high income countries. According to WHO, the annual number of ARI-related deaths in children less than five years old (excluding death caused by measles, pertussis and neonatal deaths) were about 2.1 million i.e., about 20% of all childhood deaths. Each year about 10.8 million children die because of ARI. Estimates indicate that in 2000, 1.9 million children died due to ARI, 70% of them in Africa and Southeast Asia (Williams *et al.,* 2002).

Acute respiratory infections (ARIs), principally pneumonia, account for approximately 1.9 million (1.6 - 2.2 million) deaths globally in children under 5 years of age each year, 90% of which occur in the developing world (Montasser *et al.,* 2012). In low and middle income countries, acute respiratory infections (ARIs) are a major cause of morbidity and mortality among children less than five years of age. Indeed, it is estimated that about 126 - 156 million cases of acute lower respiratory infections (ALRI) such as pneumonia and bronchiolitis occur worldwide each year in children leading to approximately 1.4 million deaths. More than 95% of these deaths occur in Africa and in South-East Asia (Liu L.*et al.*, 2012, Nair H.*et al.*, 2013, Sonego M.*et al.*, 2015).

Every year ARIs account for over 12 million hospital admissions among children below five years of age (Nair H.*et al.*, 2013). Nine out of ten children living in extreme poverty were from Sub-Saharan Africa, this region is the home of the majority of the worlds pre-school age children. Compared to children from richer households, children from poor households are more likely to be exposed to poor environmental conditions, such as indoor air pollution due to cooking fuels that may exacerbate childhood diseases, specifically ARI symptoms (UNICEF, 2017, IVAC, 2016).

Study conducted in Bangladesh reported that 21.3% under five children suffered from ARI two weeks preceding the survey and the Prevalence of severe ARI is higher among children born to mothers with primary education compared to children of mothers completing secondary or higher education (Azad, 2009). Malnourished children from a lower socioeconomic category are more likely to suffer from ARI (Geberetsadik *et al.*, 2015). According to the 2016 Ethiopia Demographic and Health Survey (EDHS),the prevalence of ARI was 7% (CSA, 2016) and other study conducted in the Northwest, Ethiopia showed overall two week prevalence of acute respiratory infection and pneumonia among under-fives was 20.6% and 16.1% respectively(Fekadu, 2014).

According to 2012 central statistical agency report, there is the highest burden of pneumonia in Ethiopia that is 88 in 1,000 children under age 5 die before their fifth birthday (CSA, 2012). Acute respiratory infection (ARI), and particularly pneumonia, accounts for 18% of death in Ethiopia; improving early care are a key strategy for early diagnosis and treatment (UNICEF, 2014). Integrated management of common childhood illness and community case management are among the program initiatives scaled up nationally to address ARI (Miller, 2014).

The 2016 Ethiopian Demographic and Health Survey (EDHS) data used for this study are based on two stage stratified cluster sampling. The appropriate approach to analyzing ARI among under five children from this survey is therefore based on nested sources of variability. Here the units at lower level (level-1) are individuals (under-five children) who are nested within units at higher level (region). Beside the nested source of variability; the response variable in this study is ARI status of under five children which is binary response. Because of this, multilevel logistic regression analysis considers the variations due to hierarchy structure for binary response. Therefore, for analyzing the data, Bayesian multilevel approach was used.

With this study, the reason why the Bayesian approach is preferred over the usual frequentist technique is that the power of information obtained from the approach is much better as it is the combination of likelihood data and prior information about the distribution of the parameter. The other advantage of Bayesian approach is the possibility of improving the precision of the results by introducing external information in terms of the priori distribution (M.L. Calle *et al.*, 2006). Various studies show that Bayesian approach may have advantages over the frequentist one, particularly in case of a low power of the frequentist analysis (Salameh, 2014).

Most studies in this area was tried to use statistical models like Binary logistic regression model to identify the determinant factors of ARI. Those studies were more of classical, such studies have no power to handle the geographical variation as well as the variation within individuals. Due to this they didnt focus on regional variations. In another way, classical model is naturally less accurate than the Bayesian approach models since all information in the classical has been obtained from likelihood only wherein the Bayesian approach it was the integration of prior information and that of likelihood. Thus, to fill the gaps with those previous studies we investigate the factors associated with ARI among under-five children and regional variations using, Bayesian multilevel logistic regression model. In this study, we present the Bayesian estimation of hierarchical models, using Markov Chain Monte Carlo (MCMC) simulation methods (Gilks, 1995, Browne *et al.*, 1998).

ARI cannot be tackled without understanding its causes; there is also inconsistency across studies regarding the determinants factor behind occurrence of ARI in under-five children therefore that is why this study is crucial to assess the prevalence of ARI and identify underling factors of ARI among under-five children in Ethiopia. Therefore, the main aim of this study was to investigate the determinant factors associated with acute respiratory infections among under five children in Ethiopia, using Bayesian multilevel logistic regression model.

1.2 Statement of the Problem

In 2016, an estimated 5.6 million children died before reaching their fifth birthday (UNICEF. 2016).This due to lack of health care services, poverty and low levels of education in developing countries are all major risk factors that are contributing to high mortality rates of ARI in under-five children. Acute Respiratory Infections (ARI) like pneumonia are among the leading causes of morbidity and mortality in children under five in developing nations (Hug L., 2017). Children living in developing countries are ten to fifty times more likely to die from ARI than those living in developed countries (Chatterjee S., 2016). In Ethiopia, about 190, 000 children are still dying each year, although Ethiopia has achieved fourth MDG target three years earlier by reducing under-five mortality by 67% from the 1990 estimate. Based on the 2014 WHO estimates ARI is the leading causes of under-five mortality, which accounts 18% of total death among under five children (FDRE, MOH, 2015). Also a study conducted in Ethiopia by Integrated Community Case Management Survey in Amhara, SNNP, and Tigray Regions revealed that 19% of the children had had ARI in the preceding two weeks (A survey of the last ten kilometers project, 2013).

The achievement of health care intervention depends on a correct understanding of the demographic, socioeconomic, health, environmental and nutritional factors which may influence the prevalence of disease and deaths. So far, there are no such detailed studies conducted to explore all aspects of the prevalence of ARI among under five children in Ethiopia, particularly the effects of socioeconomic factors, regional variation, and factors that contribute for a regional variation on the occurrence of ARI among under five children using Bayesian multilevel logistic regression model.

Several researchers such as (Acquah, 2013) studied the comparison of Bayesian and classical and found that Bayesian gave a better result than the classical statistics. Other studies have also shown similar result (Kawo *et al.*, 2018). In the frequentist approach a parameter is treated as a single fixed value or constant, in the Bayesian approach a parameter is treated as a random variable with a specific parameter distribution (Gelman *et al.*, 2014). The Bayesian approach gives the possibility of incorporating additional information that is external to the sample by prior distributions (Gelman

A., 2000, Ibrahim J.G., 2000). This additional information may improve accuracy and credibility of estimations. The credible regions incorporate this prior information, while frequentist confidence intervals are based only on the sample data.

Some scholars have used logistic regression to identify important risk factors of acute respiratory infections (ARIs) among under-five children (Geberetsadik *et al.*, 2015, Harerimana *et al.*, 2016), which actually cannot be empowered to answer whether there were geographical variations or not. i.e. these statistical methodologies are not capable to consider the regional variations. Although, classical technique fits the logistic regression by means of an iterative approach and in some cases, as a result of this iterative approach, convergence may be difficult to achieve. The robustness and accuracy of the results produced by Bayesian approach make its gain popularity in data analysis. Therefore, considering the seriousness of the disease and gaps found with different studies, this study was intended to fill the gap on this regards by considering the random effects under the multilevel model of the Bayesian approach because multilevel consider the variations due to the hierarchical structure in the data. Also it allows the simultaneous examination of the effects of group level (regional) and individual level variables. Other studies conducted at hospital level with very limited covariates; that have relative drawback. Thus, this study mainly concentrated on Bayesian multilevel approaches to identify significant factors associated with ARI among under five children.Therefore, this study aims to address the following research questions:

1. Are socioeconomic, demographic,health, environmental and nutritional characteristics experiencing ARI?

2. Do regions differ in the occurrence of ARI among under-five children?

3. From the study variables which predictors have variation across regions?

4. Which model is a good fit from empty model (null model), intercept model and random coefficient model?

1.3 Objectives of the study

1.3.1 General Objective

The general objective of this study was to investigate factors associated with the prevalence of ARI morbidity among under five children in Ethiopia: using Bayesian multilevel logistic regression model.

1.3.2 Specific objectives

- To assess the effect of socioeconomic, demographic, health, environmental and nutritional related factors associated with occurrence of ARI among under-five children.
- To investigate between and within regional variations of ARI among under-five children in Ethiopia.
- To determine from the study variables, the variation of predictors across regions.
- To compare empty model (null model), the intercept model and random coefficient model.

1.4 Significance of the study

The following are few points that can be taken as significance of this study to the users:

- The result of this study provide information to government and other concerned bodies to make enabling environment for the intervention to reduce the occurrence of ARIs.
- It helps to reduce the occurrence of ARI among under-five children by giving awareness for the society on the factors that increase the probability of a child suffered from ARIs.
- To give emphases on the factors that have strong association with the occurrence of ARI among under-five children.

- This study also useful to understand how important it is to consider the hierarchical structure of the data, whether the magnitude of the random effects is small or large. It is specifically helpful for those who want to deal with the variation between and within the clusters or groups for cross sectional data set of the factors that affect Respiratory systems of under-five childrens.
- It serves as stepping-stone for those who are interested to undertake in depth research on issues related to the incidence of ARI.

1.5 Scope of the study

This study uses secondary data collected by the Central Statistical Authority (CSA), i.e., the Ethiopia Demographic and Health Survey of the year 2016 (EDHS 2016). The study uses information on women of age 15-49 who had at least one child in the five years before the survey. The study is used to identify the risk factors of ARI among under five children using the Bayesian multilevel logistic model.

1.6 Organization of the Study

This study was structured into five sections. The first chapter gives a general background of the study, statement of the problem, objective, and its significance of the study. Chapter 2 deals with the review of literature on ARI among under five children in Ethiopia and the rest of the world, whereas chapter three specifies the data and methodology of the study such as sources of data and variables to be included in the study. Methods of data analysis are also described in this chapter. Chapter 4 reports results from the Bayesian multilevel logistic regression analysis and provides discussions. Finally, the last chapter draws conclusions and makes recommendations for further studies.

2 Literature Review

2.1 Overview of Acute Respiratory Infection Cases

In 2015, the global under-five mortality was 43 per 1000 live births, a 44% decline as compared to the rates in 2000 (WHO, 2017). Newborn deaths were about half of all under-five deaths in all WHO regions, except for WHO African Region where a third of under-five deaths happened after the first month of life (WHO, 2017). The global under-five mortality rate in 2015 was 43 per 1000 live births, while the neonatal mortality rate was 19 per 1000 live births representing declines of 44% and 37% respectively compared to the rates in 2000. Estimates also indicate that 30-50% of outpatient department attendance and 20-40% of hospital admissions may be attributed to acute respiratory infections and pneumonia (Walke *et al.,* 2014).

Acute respiratory tract infections (ARIs) are classified as upper respiratory infections or lower respiratory tract infections (Simoes *et al.*, 2006). The upper respiratory tract comprised of the airways from the nostrils to the vocal cords in the larynx, and include the paranasal sinuses and the middle ear. The continuation of the airways from the trachea and bronchi down the bronchioles and alveoli are regarded as the lower respiratory tract (Simoes *et al.*,2006).

In the developing world, ARI represent the major causes of mortality and morbidity among children under five years (Ujunwa and Ezeonu, 2014). Children under five years are high risk group and vulnerable to respiratory infection as this is the developmental stage of their physical growth and lung function (Liu *et al.,* 2013). Their respiratory system is not completely developed until the age of six. Compared to adults, children breath more air in proportion to their body weights. The incidence of ARI in children for the first five years of life is about 6-8 episodes (Ujunwa and Ezeonu, 2014).

2.2 Determinants of ARI in under-five children

Risk factors for ARI can be categorized in three levels: on the societal level, there are risk factors such as socioeconomic level of a community; on the household level factors such as parental characteristics, household size and income, and indoor air pollution; and on the individual level characteristics of children such as the presence of malnutrition (Harerimana *et al.,* 2016).

2.2.1 Socio - economic and Demographic factors

Various studies identified different socioeconomic and demographic variables as influential for the occurrence of acute respiratory infection such as household wealth index, mathers educational level, child age, maternal age and maternal working status.

The study conducted in Pakistan to establish health care seeking patterns for childhood illnesses revealed that, mothers educational status was strongly associated with type of provider sought for childhood illness (p = 0.001). In addition, family income and socio-economic status of the household were also associated with health care seeking (Rehman, 2014). Also, another study carried out in Sub Saharan Africa, it revealed that caregivers with high education level and those from high wealth quintiles were more likely to seek appropriate care for their children (Noordam *et al.,* 2015).

The study done in India showed that low socioeconomic status (OR=4.89), were found to be significant risk factors of ARI (Bhat, 2013). Also, children within the age group 6-23 months were more likely to have experienced the symptoms of ARI compared to children below 6 months of age or older children (Mishira, 2013).

A lancet systematic analysis also examined that ALRI incidence was highest in neonates aged 0-27 days and infants aged 0-11 months (Nair H, 2013). A study done in Bangladesh using logistic regression revealed that age of child has a significant association with incidence of ARI. This variable had three groups, infant (age 12 months), toddler (age 12-23 months), and child (24-59 months). It was found that an infant had 1.8 times and a toddler had 1.5 times significantly higher odds of suffering from ARI than a child (Azad, 2009). Also the study done in Rwanda revealed that families having more than two under five children at home were significantly associated with ALRI and ALRI was particularly higher among children less than two years (0-11 months: 5.2%; 12-23 months: 5.1%) (Harerimana *et al.,* 2016).

A longitudinal cohort study conducted on 400 children in Gulbarga to assess the morbidity pattern and determinants of ARI found a significantly higher susceptibility to ARI of boys than girls (Ramani *et al.,* 2016).The likely reason highlighted by this study could be that boys spend more time outdoors than girls thereby increasing their risk of exposure. The study done in Kenya revealed that, children living in urban areas are 1.674 more times likely to develop ARI as to those of the same age living in rural areas (Muthoni *et al.,* 2017).

The study conducted to determine healthcare-seeking behavior for childhood illnesses in Kenya revealed that, children from households with higher socioeconomic status were more likely to be taken to a hospital than those children from low socioeconomic status (Burton *et al.,* 2011). Likewise, another study carried out in Nandi County, it was shown that, mothers with higher levels of education were more likely to seek immediate health facility care compared to those with lower levels of education (Keter, 2015).

The study conducted in Kenya using Bayesian hierarchical spatial modelling approach to investigate the spatial variation in the prevalence of ARI revealed that, both geographical heterogeneity and the high prevalence rate of ARI was observed. Accordingly, disease prevalence is often associated with many socioeconomic status factors such as overcrowding, unemployment rates, educational and housing quality were significantly affected under five children. From the analysis the risk of ARI was found to vary in a decreasing manner from poorest, poorer and middle class with the highest being in the household in the poorest category. However, there was a reduction in the risk in children whose households were richer. But, mothers age, and mother's education level were found to be insignificant (Muthoni *et al.,* 2017).

Geberetsadik *et al.,* (2015) revealed that based on logistic regression analysis, the odds of ARI decreased as the age of the child increased. A 50% reduction in the odds of ARI was followed in children aged 48-59 months as compared with infants under 6 months of age. Study conducted in the Wondo Genet district, Sidama zone, SNNPR, Ethiopia showed that children at the age range of 2-12 months were 4 times more likely to develop pneumonia as compared to older age groups (Abuka T., 2016).

Ethiopia is a diverse country and incidence of ARI in under five children is not evenly distributed throughout the country. Regional disparities in prevalence of ARI among under five children are associated with factors at the community level that distinguish these regions from each other. The availability of services and social amenities in communities, or the lack of infrastructure, may positively or negatively influence the health of the residents of communities. Some of these factors include differences in community-level development, population density, the prevalence of poverty, and the availability of child health care services. The study developed on Acute Respiratory Infection Symptoms among under Five Children in Ethiopia by using multilevel analysis shows the odds of having ARI symptoms estimated to be higher among children in Tigray, Amhara, Oromia, Somale, Benishangul-Gumuz and Gambela regions compared to those in Dire Dawa. From regional and two city administrations of Ethiopia, Children in Dire Dara had the lowest exposure to ARI symptoms. As this study shows both mothers education level, and mothers occupation emerged among the most important predictors of incidence of ARI among under five children (Jabessa S, 2015).

The study carried out in Ethiopia to establish healthcare-seeking behavior for ARI, it revealed that children from the two highest economic quintiles were 35 to 39% more likely to seek health care immediately as compared with households in lower economic quintiles. In addition, household heads with informal education were 1.6 times more likely to take their children to a health facility compared with household heads with no education households (Mebratie *et al.,* 2014). Also other study revealed that, the odds of under five children having mother with no education, to have symptoms of ARI is increased by 16.69% as compared to above the secondary educational level and the risk of ARI symptoms among under five children is significantly less, on average, for children whose mothers are not working compared to whose mothers have work (Jabessa S, 2015).

2.2.2 Health, nutritional and environmental factors

Several studies show that there are many nutritional, health and environmental factors like types of cooking fuel, duration of breastfeeding, received vitamin A recently, stunting, wasting, history of diarrhea and source of drinking water is important determinant factors for childhood health status related to ARI.

Environmental factors similarly increases a childs susceptibility to ARI: such as indoor air pollution caused by cooking and heating with biomass fuels (such as wood or dung); living in crowded homes; and parental smoking (WHO, 2016). The risk of

ARI, ALRI in children and exposure biomass fuels in the world especially in developing countries. According to this study, exposure with biomass fuels was recognized as a cause of lower respiratory infection in children. In need, resulting increased risk, lower respiratory infection and acute respiratory infection in children (Mohammadi *et al.,* 2018).

Environmental conditions such as contaminated water and inadequate sanitation can contribute to acute diarrhea while poor air quality can be a factor to ARI. Multiple exposures to environmental contaminants may increase the occurrence of acute diarrhea and ARI among children under five years of age (Briggs, 2003). Acute diarrhea can lead to acute weight loss, malnutrition and stunting, which are risk factors for ARI in a low income setting. This association has a considerable public health importance (Schmidt, 2009). Problem of both diseases can be duly reduced by implementing community-based prevention strategies such as improved water quality, sanitation and hygiene, and better quality of fuel for cooking (Ghimire M, 2010).

The WHO defines exclusive breastfeeding as the practice of only giving an infant breast-milk for the first six months of life (WHO, 2018). A systematic review of 19 studies that investigated 19 risk factors of acute lower respiratory tract infection found the lack of exclusive breastfeeding (OR = 2.34) as one of the seven risk factors that were significantly associated with severe acute lower respiratory tract infection (Jackson *et al.,* 2013).Other study done in Netherlands has examined that compared with never-breastfed infants, those who were breastfed exclusively until the age of 4 months and partially thereafter had lower risks of infections in LRTI until the age of 6 months (OR = 0.50) and of LRTI infections between the ages of 7 and 12 months (OR = 0.46) (Taylor, 2012).

Azad, (2009) showed that a child without taking vitamin A in the last six months had 29% higher odds of suffering from ARI. Children from poorest families had 30% higher odds of suffering from ARI. A child with stunted growth has 19% more odds of suffering from ARI ($p < 0.001$) compared to children without stunting growth.

Bbaale, (2011) revealed that child age and nutritional status are imperatively associated with the incidence of diseases. Children who are 0-12 months of age, children in older age cohorts reduce the probability of occurrence of diarrhea and ARI by 16-23%

and 11-20%. Another study done using case control study showed that childrens having, history of diarrhea (AOR = 3.06) and household history of acute lower respiratory infection (AOR = 3.04) respectively, were at higher odds of developing pneumonia (Dadi *et al.,* 2014). In Nepali children, diarrhea and acute LRI occurred simultaneously more than chance alone. Incidence of ALRI increased as the number of days with diarrhea in the prior 28 days increased; the greatest incident rate ratio was reported among children with 20 or more days of diarrhea (OR=1.02) in Nepal (Walkeret *et al.,*2013).

Ekure *et al.,* (2013) a study carried out in Nigeria to establish level of mothers knowledge on childhood pneumonia showed that about 50% were able to correctly identify fast or difficult breathing as a symptom of pneumonia. However, 75% listed cold as a cause of childhood pneumonia. Furthermore, 75.8% and 49.5% of the mothers respectively said reducing exposure to cold and wearing warm clothes were the two commonest way of preventing pneumonia.

Geberetsadik *et al.,* (2015) revealed that the prevalence of ARI was found to be higher in children with malnutrition. The odds of ARI among severely wasted children were 1.7 times higher than in normal children. Also the study shows that age of the child, fathers educational status, and maternal occupation were statistically significant. But, duration of breastfeeding, had received vitamin A in the previous 6 months and place of residence were statistically insignificant.

2.3 Overview of Bayesian Multilevel Modeling

The Bayesian approach presents a well-established framework for making an inference from observed data for quantities of interest by using an underlying probability model for a comprehensive overview of modern Bayesian statistical analysis. The Bayesian methodology differs from the classical frequentist approach in that all of the unknown parameters in the underlying probability model are treated as random variables, as opposed to unknown constants in the classical frequentist approach. As such, the unknown parameters are assigned prior distributions which are based on a priori subjective beliefs or scientific knowledge about the unknown parameters. In other words, prior distributions serve as probabilistic descriptions of what is known about the unknown parameters before observational data are collected and analyzed (Gelman et al., 2013; Berger, 2013).Bayesian methods may also improve on classical estimators in terms of the precision of estimates. This happens because of specifying the prior that brings extra information or data based on accumulated knowledge, and the posterior estimate is based on the combined sources of information (prior and likelihood) therefore has greater precision (Richardson, 2003).

The study conducted using classical and Bayesian approach in the logistic regression model on the child's attention deficit hyperactivity disorder revealed that the Bayesian approach in statistical analysis is an alternative to be considered, given that it makes it possible to introduce prior information about the phenomenon under study. The result shows a reduction of standard errors associated to the coefficients obtained from the Bayesian analysis, thus bringing a greater stability to the coefficients. Therefore, they concluded that Bayesian methods provide a more precise and powerful result (Gordvil-Merino *et al.,* 2010).

The study applied to longitudinal data on pregnancy complication in rural Bangladesh. As the authors comparisons between classical and Bayesian approach. In the Bayesian approach under squared error of loss function were used and compare with method of maximum likelihood and we have found that length of the Bayesian credible interval is smaller than length of the confidence interval for classical approach. So, Bayesian estimation found to be better.(Mahanta *et al.,* 2015).

The study developed on anemia prevalence among children aged 6-59 months in Ethiopia by comparing classical and Bayesian approaches. As the author comparison between Bayesian approach and classical approach results indicated a reduction of standard errors is associated with the coefficients obtained from the Bayesian approach and Bayesian produces precise estimates and more robust compared to the classical (Kawo *et al.,* 2018).

Multilevel logistic regression is a powerful statistical modeling method that allows the incorporation of explanatory variables at different levels of hierarchy so that under five children and regional level clustering of the outcome are taken into account. Generally ignoring the correlated or nested data can completely be resulted with the wrong estimation which in turn leads to a wrong conclusion (Sainani, 2010). Therefore, the nature of EDHS data is hierarchical in which individuals are nested within regions for which multilevel models are advisable. Moreover, the Bayesian approach offers a natural solution to the problem of multiple comparisons, when the situation is adequately modeled in a multilevel framework (Gelman, 2010), and allows a priori knowledge to be incorporated in data analysis via the prior distribution.

A number of efficient algorithms are available for obtaining maximum likelihood (ML) estimates of a multilevel model, for example the iterative generalized least squares procedure (IGLS) or restricted maximum likelihood estimates (RIGLS). Nevertheless, Bayesian methods can implement multilevel models without statistical limitations. Bayesian MCMC methods yield inferences based upon samples from the full posterior distribution and allow exact inference in cases where, as mentioned above, the likelihood based methods yield approximations. Here, we apply a fully Bayesian approach as suggested in (Fahrmeir L, 2001) which is based on Markov priors and uses Markov Chain Monte Carlo (MCMC) techniques for inference and model checking. There is no established method for determining an appropriate number of iterations and burn-in size. Rather, the researcher use a trial-and-error process in which the ultimate goal is to obtain stable parameter estimates that minimize simulation error. As with the computational intensity this steps require more time on the part of the researcher. However,for choice of the model, we routinely used the DIC developed in (Spiegelhalter *et al.,* 2002), as a measure of fit and model complexity.

3 Data and Methodology

3.1 Description of the Study Area

Ethiopia is situated in the Horn of Africa between 3 and 15 degrees north latitude and 33 and 48 degrees east longitude. It is the second-most populous nation in Africa next to Nigeria. Ethiopia is bordered by Eritrea to the North, Djibouti, and Somalia to the East, Sudan and South Sudan to the West, and Kenya to the South. Ethiopia has eleven geographic or administrative regions: nine regional states (Tigray, Afar, Amhara, Oromia, Somali, Benishangul-Gumuz, SNNPR, and Harari) and two city administrations (Addis Ababa and Dire Dawa) with a capital city of Addis Ababa.

3.2 Source of data

The dataset in this study was obtained from the Demographic and Health Survey conducted in Ethiopia in 2016. The 2016 Ethiopia Demographic and Health Survey (EDHS) was the fourth survey conducted in Ethiopia as part of the worldwide Demographic and Health Surveys project. It was implemented by the Central Statistical Agency (CSA) at the request of the Federal Ministry of Health (FMoH). Data collection took place from January 18, 2016, to June 27, 2016 with national representative of 18,008 households were selected based on a nationally representative sample that provides estimates at the national and regional levels and for urban and rural areas. The data provide in-depth information on family planning, fertility, marriage, infant, child, adult and maternal mortality, maternal and child health, gender, nutrition, malaria, knowledge of HIV/AIDS and other sexually transmitted diseases.

The 2016 EDHS sample was selected by considering two-stage cluster design and census enumeration areas (EAs) were the sampling units for the first stage. A typical two-level stratification involves first stratifying the population by region at the first level and then by urban-rural within each region. The sample included 645 EAs (202 in urban areas and 443 in rural areas).In the sampling procedure, households comprised the second stage of sampling. A complete listing of households was carried out in each of the 645 selected enumeration areas by equal probability systematic sampling according to proportional to EAs measure of size from January 18, 2016, to June 27, 2016.

All women aged 15-49 years who had at least one child in the five years before the survey were eligible for participation. To handle missing values we used list wise deletion which is a common approach and easy to perform by deleting all incomplete observations from the analysis. The result was unbiased when data are MCAR. Even so, the disadvantage of this method is reduction of sample size. The sample for this study would be consisted of 10,641 under-five children, from which only 9,918 of them would be considered in this study.Thus, the analysis for this study was on the occurrence of acute respiratory infection would be presented based on children less than five years age.

3.3 Variables in the Study

Variables considered in this study were selected based on literature review such as socio-economic, demographic, health, environmental and nutritional related factors are considered as the risk factor for the occurrence of ARI in under-five children.

3.3.1 The response variable

In this study the response variable was the presence or absence of ARI symptoms in a child who was under five years of age, which was coded with a value zero to indicate absence of ARI symptoms and one to indicate presence of ARI symptoms. The two conditions required for a child to be classified as having ARI were having a cough and short, rapid breaths in the last two weeks, which were measured based on mothers reports about the symptoms of these conditions. The symptoms are compatible with ARIs. Therefore the i^{th} child in the j^{th} region is represented by a random variable Y_{ij}, with two possible values coded as 1 and 0. Hence, the response variable for the i^{th} child in the j^{th} region is measured as a dichotomous variable.

$$Y_{ij} = \begin{cases} 1, \text{ if the } i^{th} \text{ child suffered from ARI in the } j^{th} \text{ region} \\ 0, \quad \text{otherewise} \end{cases} \tag{1}$$

With i=1,2,3,....n and j=1,2,3,....,k. Where: n- is the number of under five children in each region j and k-is the number of regions.

Let denote the proportion of success (child suffered from ARI)

$$P(Y_{ij}=1)=\pi_{ij}, P(Y_{ij}=0)=1-\pi_{ij} \tag{2}$$

and $Y_i \sim Bernoulli(\pi_i)$

3.3.2 Explanatory Variables

The independent variables that are used in this study will be classified as socio-economic, demographic, health, environmental and nutritional related factors.

Demographic and socio-economic characteristics:- In this study maternal education level, mother working status, household wealth index, place of residence,sex of child, child age, number of children, maternal age and geographical region would be expected as demographic and socio-economic risk factors.

Health, environmental and nutritional related characteristics:- There are certain health, environmental and nutritional related characteristics that may affect the occurrence of ARI among under-five children were: received vitamin A recently, wasting (acute malnutrition), stunting (chronic malnutrition), duration of breast feeding, fuel used for cooking, source of drinking water and history of diarrhea are important health, environmental and nutritional characteristics that had been included in this study.

Table 3.1: Description of Demographic and socio-economic factors and categories.

variable names	codes and categories used in this study
Maternal education	(0) No education (1) Primary (2) Secondary and above
Mother currently working	(0) Not working (1)Working
Place of residence	(0) Rural (1) Urban

Household wealth index	(0) Poorest
	(1) Poorer
	(2) Medium
	(3) Richer
	(4) Richest
Current age of the child	(0) less than 6 months
	(1) 6-11 months
	(2) 12-23 months
	(3) 24-35 months
	(4) 36-47 months
	(5) 48-59 months
Region	(1) Tigray
	(2) Afar
	(3) Amhara
	(4) Oromia
	(5) Somali
	(6) Benshangulgumize
	(7) SNNP
	(8) Gambela
	(9) Harari
	(10) Addis Ababa
	(11) Dire Dawa
Sex of the child	(1) Male
	(2) Female
Number of children	(0) 1-3 child
	(1) 4-6 child
	(2) above 6 child
Maternal age	(0) 15-19
	(1) 20-34
	(2) 35-49

Table 3.2: Description of Health, environmental and nutritional related characteristics and categories.

variable names	codes and categories used in this study
Received vitamin A recently	(0) No
	(1) Yes
Wasting	(0) Not wasted
	(1) Wasted
Stunting	(0) Not stunted
	(1) Stunted
Duration of breast Feeding	(0) Never breast feeding
	(1) Ever breast feeding, not currently
	(2) Still breast feeding
Types of cooking fuel	(1) unsafe/unclean
	(2) safe/clean
Source of drinking water	(0) Unprotected
	(1) Protected
Had diarrhea	(1) Had no diarrhea
	(2) Had diarrhea

3.4 Methods of Data Analysis

The statistical analysis in this research is based on Bayesian multilevel logistic model. The data collection procedure is the hierarchical level or structures that means the levels are nested one another; thus why the reason for selecting this model. The data analysis was taken place using MLwiN 2.02 and SPSS version 20 software.

3.4.1 Bayesian multilevel logistic regression model

In this study a Bayesian multilevel logistic regression approach for binary outcomes was preferred, which takes into account the hierarchical structure of data and properly estimates the parameters and accuracy intervals (Bornmann *et al.*, 2013).The multilevel logistic regression model is appropriate for research designs where data for respondents are organized more than one level i.e, nested data. The units of analysis are individuals at a lower level i.e, under five-children in our case who are nested within aggregate units at a higher level i.e. regions. A multilevel logistic regression model is also referred to as a hierarchical logistic regression model, or as random effects (mixed effects) logistic regression model. The multilevel logistic regression extends from single level logistic regression model by including random effects to the model (Snijders and Bosker, 1999).

The appropriate approach to analyzing under-five childrens data from this survey is therefore based on nested sources of variability. Here the units at lower level are under-five children who are nested within units at higher level (regions). Due to this nested structure, the odds of childs experiencing ARIs are not independent, because children from the same cluster (region) may share common exposure to the outcome of interest. The response variable in this study is ARIs status which is binary and hence multilevel logistic regression model is a natural choice for modeling. In this data structure, level-1 is the under-five childrens level and level-2 is the regional level. Within each level-2 unit, there is n_j in the j^{th} region.

We further simplify the presentation by assuming there is under-five children-level predictor and regional level factor of ARI. To provide a familiar starting point, we will first consider a two-level model for binary Outcomes with a single explanatory variable. Suppose we have data consisting of under-five childrens, (level one) grouped into regions (level two).Let Y_{ij} be the binary response for ARI among i^{th} under five children in region j and X_{ij} be an explanatory variable at the under five children level.We define the probability of the response equal to one $\pi_{ij} = P(y_{ij} = 1)$ Where;π_{ij} be modeled using a logit link function. The standard assumption is that Y_{ij} has a Bernoulli distribution. Then, the two-level models are given by:

$$logit(\pi_{ij}) = log[\frac{\pi_{ij}}{1-\pi_{ij}}] = \beta_{oj} + \Sigma_{h=1}^{k}\beta_{hj}X_{hij} \tag{3}$$

where i=1,2....n_j,h=1,2...k and j=1,2,....11

$$\beta_{oj} = \beta_o + U_oj, \beta_{1j} = \beta_1 + U_{1j}...\beta_{kj} = \beta_k + U_{kj}$$

$$Logit(\pi_{ij}) = Log[\frac{\pi_{ij}}{1-\pi_{ij}}] = \beta_o + \Sigma_{h=1}^{k}\beta_{hj}X_{hij} + U_{oj} + \Sigma_{h=1}^{k}U_{hj}X_{hij} \quad (4)$$

$X_i = (X_{1ij}, X_{2ij}....X_{kij})$ represent the first and the second level covariates, for variable k ($\beta = \beta_o, \beta_1, ...\beta_k$) are the regression parameter coefficient. $U_{oj}, U_{1j}, ..., U_{kj}$ is the random effect of the model parameter at level two. With the assumption U_{uj},follows a normal distribution with mean zero and variance σ_u^2. Without U_{hj} the above equation can be the single-level logistic regression. That means the 3^{rd} equation is the single level logistic model and the 4^{th} equation is two levels model. Therefore conditional on $U_{oj}, U_{1j}, ..., U_{kj}$ the Y_{ij} can be assumed to be independently distributed as Bernoulli random variables.

3.4.2 Multilevel Analysis of Empty Model (Null Model)

The empty two-level model for a binary outcome variable refers to a population of groups (level-two units (regions)) and specifies the probability distribution for group-dependent probabilities P_j in $Y_{ij} = P_j + \epsilon_{ij}$ without taking further explanatory variables into account. Here, the logit transformed model,$logit(\pi_{ij})$ can have the normal distribution. Consequently, the empty model can possibly be expressed in the form of the following formula:

$$Logit(\pi_{ij}) = \beta_o + U_{oj} \quad (5)$$

In the equation above, β_o indicates the population average of the transformed probability and U_{0j} is the random deviations from this average for region j. The residual term that is associated with the group dependent deviations, U_{0j} has a unique effect of region j on the response variable; and it is assumed to be normally and independently distributed with mean zero and variance ,σ_o^2 that is $U_{oj} \sim N(0, \sigma_0^2)$.In this situations, the level 2 residual can possibly capture the variations across regional means. In this model, the amount of variance regarding ARIs that is attributable within group characteristics (under-five children) and between group difference (region) can be investigated.

Equation (5) does not include a separate parameter for the level one variance (Snijders, 1999).The reason is the level one residual variance of binary outcome variable follows directly the success probability indicated as follow: $Var(\epsilon_{ij}) = \pi_{ij}(1-\pi_{ij})$ where ϵ_{ij} is a child suffered from ARIs dependent residuals.

The other reason of applying multilevel analysis is the existence of intra-class (intra-regional) correlation arising from similarity of under-five children in the same region compared to those of different regions.

The intra-class correlation coefficient (ICC) measures the proportion of variance in the outcome explained by the grouping structure. ICC can be calculated using an intercept-only model or an empty model. The ICC can be calculated as:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} \tag{6}$$

Where σ_u^2 is the variance between the group which can be estimated by U_{oj} and σ_e^2 is within-group variance (John, 2009). Denote π_o the probability corresponding to the average values β_o as defined by $P(\pi_o) = \beta_o$ for the logit function, the so-called logistic transformation of β_o, is defined as:

$$\pi_o = logit(\beta_o) = \frac{e^{\beta_o}}{1+e^{\beta_o}} \tag{7}$$

Note that due to the non-linear nature of the logit link function, there is no simple relation between the variance of the deviations U_{oj} . However, there is an approximate formula which is valid when the variances are small and is given by:

$$Var(\pi_j) = (\pi_o(1-\pi_o))^2\sigma_o^2 \tag{8}$$

Note that an estimate of population variance $Var(\pi_j)$ can be obtained by replacing sample estimates of π_o and σ_o^2. The resulting approximation can be compared with the nonparametric estimate:

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

Hypothesis:
H_0: There is no regional variation in ARI among under-five children in Ethiopia.
H_1: There is a regional variation in ARI among under-five children in Ethiopia.

3.4.3 Bayesian Multilevel Analysis of Random Intercept Model

The random intercept model is used to model unobserved heterogeneity in the overall response by introducing random effects. In the random intercept model the intercept is the only random effect meaning that the groups differ with respect to the average value of the response variable, but the relation between explanatory and response variables cannot be differ between groups. The random intercept model expresses the log odds, i.e the logit of π_{ij} , as a sum of linear functions of the explanatory variables. That is,

$$Logit(\pi_{ij}) = log[\frac{\pi_{ij}}{1-\pi_{ij}}] = \beta_{oj} + \Sigma_{h=1}^{k}\beta_h X_{hij} \tag{9}$$

i=1,2,.....,n, j=1,2,....,11
Where the intercept term β_{0j} is assumed to vary randomly and is given by the sum of an average intercept β_0 and group-dependent deviations; $\beta_{0j} = \beta_0 + U_{0j}$.
As a result we have:

$$Logit(\pi_{ij}) = \beta_o + \Sigma_{h=1}^{k}\beta_h X_{hij} + U_{oj} \tag{10}$$

Solving for π_{ij}

$$\pi_{ij} = \frac{e^{\beta_o + \Sigma_{h=1}^{k} X_{hij} + U_{oj}}}{1 + e^{\beta_o + \Sigma_{h=1}^{k}\beta_h X_{hij} + U_{oj}}} \tag{11}$$

Equation (10) does not include a level one residual because it is an equation for the probability π_{ij} rather than for the outcome Y_{ij}, where $\beta_o + \Sigma_{h=1}^{k}\beta_h X_{hij}$ is the fixed part of the model. The remaining U_{0j} is called the random or the stochastic part of the model. It is assumed that the residual U_{0j} is mutually independent and normally distributed with mean zero and variance σ_u^2 (Snijders and Bosker, 1999).

3.4.4 Bayesian multilevel Analysis of Random Coefficients Model

Random coefficient logistic regression is based on linear models for the log-odds that include random effects for the groups or other higher level units. The random coefficients build upon the random intercept model by allowing the effects of individual predictors to vary randomly across level 2, that is, level 1 slope coefficients are allowed to take on different values in different aggregate groups. In the random coefficient model both the intercepts and slopes are allowed to differ across the regions.

Consider a model with group-specific regression of logit of the success probability $logit(\pi_{ij})$ on a single level -one explanatory variable X:

$$Logit(\pi_{ij}) = log[\frac{\pi_{ij}}{1-\pi_{ij}}] = \beta_o + \Sigma_{h=1}^{k}\beta_h X_{hij} + U_{oj} + \Sigma_{h=1}^{k} U_{hj} X_{hij} \qquad (12)$$

The term $\Sigma_{h=1}^{k} U_{hj} X_{hij}$ can be regarded as a random interaction between group and the explanatory variables. This model implies that the groups are characterized by two random effects: their intercepts and their slopes. It assumes that for different groups, the pairs of random effects $(U_o, U_h, h = 1, 2, ..., k)$ are independent and identical distributed. The random intercept variance,$var(U_{oj}) = \sigma_o^2$, the random slope variance,$var(U_{1j}) = \sigma_1^2$ and the covariance between the random effects $cov(U_{oj}, U_{1j}) = \sigma_{01}$ are called variance components (Snijders and Bosker,1999).

3.4.5 Likelihood Function

Statistical inferences are usually based on maximum likelihood estimation (MLE). MLE chooses the parameters that maximize the likelihood of the data, and is intuitively appealing. In MLE, parameters are assumed to be unknown but fixed and are estimated with some confidence. In Bayesian statistics, the uncertainty about the unknown parameters is quantified using probability. So that, the unknown parameters are regarded as random variables.

The likelihood function used in the Bayesian approach is equivalent to that of the classical inference. The joint distribution of n independent Bernoulli trials is the product of each Bernoulli densities, where the sum of independent and identically distributed Bernoulli trials has a Binomial distribution. Specifically, let $Y_{1j}, Y_{2j}, ..., Y_{ij}$ be independent Bernoulli trials with success probabilities $\pi_{1j}, \pi_{2j}, ..., \pi_{ij}$ that is $Y_{ij} = 1$(the probability of having ARI for i^{th} children in the j^{th} region) and also $1 - \pi_{ij}$ (failure probabilities) is the probability of i^{th} children not having ARI in the j^{th} region,for $i = 1, 2, ..., n$ and $j = 1, 2, ...11$. Since, the trials are independent, the joint distribution of $Y_{1j}, Y_{2j}, ..., Y_{ij}$ is the product of n Bernoulli probabilities. The probability of success in logistic regression varies from one subject to another, depending on their covariates. Thus, the likelihood function is illustrated below as product of n Bernoulli trials:

$$L(\pi_{ij}/y_{ij}) = \Pi_{ij}(\pi_{ij})^{y_{ij}}(1 - \pi_{ij})^{1-y_{ij}}$$

and the linear predictor or the logit functions is:

$$Logit(\pi_{ij}) = log[\frac{\pi_{ij}}{1 - \pi_{ij}}] = \beta_{oj} + \Sigma_{h=1}^{k}\beta_{hj}X_{hij} + U_{oj} + \Sigma_{h=1}^{k}U_{hj}X_{hij}$$

Where:

$$\pi_{ij} = \frac{exp(\beta_{oj} + U_{oj} + \Sigma_{h=1}^{k}\beta_{hj}X_{hij} + \Sigma_{h=1}^{k}U_{hj}X_{hij})}{1 + exp(\beta_{oj} + U_{oj} + \Sigma_{h=1}^{k}\beta_{hj}X_{hij} + \Sigma_{h=1}^{k}U_{hj}X_{hij})}$$

π_{ij} represents the probability of the event for subject ij who has covariate vector X_{ij} , $Y_{ij} = 1$ indicates the presence (child having ARI) and $Y_{ij} = 0$ the absence (child not having ARI) of the event for the given subject.

3.4.6 Prior distribution.

The prior distribution is a probability distribution that represents the prior information associated with the parameters of interest. It is a key aspect of a Bayesian analysis. There are two types of prior distribution: **Informative priors** and **Non-informative priors**.

An informative prior is a prior distribution that is used when information about the parameter of interest is available before the data is collected. Typically, informative prior distributions are created from historical studies, pure expert knowledge (experience) and a combination of both. Even if there is prior knowledge about what we are examining, in some cases we might prefer not to use this and let the data speak for themselves. In this study, we wish to express our prior ignorance in to the Bayesian system. This leads to non-informative priors. A non-informative prior distribution that is used to express complete ignorance of the value before the data is collected.

In this study, the prior distributions for fixed effect parameter was $P(\beta) \sim$Uniform distributions (1) and for random effect terms was, $P(1/\sigma^2) \sim Gamma(\alpha, \theta)$ where α and θ are fixed constant parameters were used.

Let us denote the parameters $\beta_o, \beta_1, ...\beta_k$ and Ω_u as prior distributions would be given as follows; $P(\beta_o) \propto 1, P(\beta_1) \propto 1, ...P(\beta_k) \propto 1$ and $P(\Omega_u) \propto inverse - wishart(m * S_u, v)$ distribution. The parameter Ω_u is the variance-covariance matrices and S_u is an estimate for the true value of Ω_u and v is the number of row in the variance-covariance matrix. The Wishart distribution is the multivariate extension of the gamma distribution; although most statisticians use the Wishart distribution in the special case of integer degrees of freedom, in which case it simplifies to a multivariate generalization of the χ^2 distribution. As the distribution χ^2 describes the sums of squares of n draws from a univariate normal distribution, the Wishart distribution represents the sums of squares (and cross-products) of n draws from a multivariate normal distribution.

3.4.7 Posterior Distribution

Posterior distribution is obtained by multiplying the prior distribution over all parameters by the full likelihood function. All Bayesian inferential conclusions are based on the posterior distribution of the model generated. Using the prior and likelihood function above the full conditional distribution of posterior parameter $\beta_o, \beta_1, ..., \beta_k$ is given by:

$$P(\beta_h \mid \Omega_u, U_{oj}, y_{ij}) \propto \Pi_{ij} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}}$$

Where $h = 1, 2, ..., k$ and the full conditional distribution of the variance-covariance parameter Ω_u has been given as:

$$P(\Omega_u/\beta_h, U_{oj}, Y_{ij}) \propto P(Y_{ij}/\beta_h, \Omega_u, U_{oj})P(U_{oj}/\Omega_u)P(\Omega_u) \tag{13}$$

Estimating β of the posterior distribution may be difficult, for this reason, we need to use the non-analytic method such as simulation techniques. The most popular method of simulation technique is Markov Chain Monte Carlo (MCMC) methods.

3.5 Estimation Techniques

3.5.1 Markov Chain Monte Carlo (MCMC) Methods

Bayesian inference is solved by randomly drawing a very large sample from the posterior distribution. The idea of drawing a large sample from the posterior distribution is called Markov Chain Monte Carlo using MCMC techniques.

The Bayesian approach applies probability theory to a model derived from substantive knowledge and theory, deal with realistically complex situations; the approach can also be termed full probability modeling. There has recently been huge progress in methods for Bayesian computation, generally exploiting modern computer power to carry out simulations known as Markov Chain Monte Carlo (MCMC) methods. The MCMC simulation is used to do the integration numerically rather than analytically by sampling from the posterior distribution of interest even when the form of that posterior has no known algebraic form (Spiegelhalter D., 2004).

3.5.2 Metropolis-Hastings Algorithm

Metropolis-Hasting algorithm is an iterative algorithm that produces a Markov chain and permits empirical estimation of posterior distributions. Therefore, in this study metropolis-hasting algorithm were used to estimate the fixed and the random effects parameters for prevalence of ARIs among under-five children. The Metropolis-Hasting algorithm (MH) generates samples from a probability distribution using full joint density function. Metropolis-Hastings algorithm correctly applied for non-Gaussian data and if the posterior distribution doesnt follow some known distribution (Gill, (2002).The metropolis-Hasting Algorithm follows the following steps:

1. Establish starting values S for the parameter:$\theta^{j=0} = S$. set $j = 1$

   The starting values can be obtained via maximum likelihood estimation.

2. Draw a candidate parameter,θ^c from a proposal density ,$\alpha(.)$.

   The simulated value is considered a candidate because is not automatically accepted as a draw from the distribution of interest. It must be evaluated for acceptance.

3. Compute the ratio $=\frac{f(\theta^c)\alpha(\theta^{j-1}|\theta^c)}{f(\theta^{j-1})\alpha(\theta^c|\theta^{j-1})}$

4. Compute R with a $U(0,1)$ random draw u. If R less than u, then set $\theta^j = \theta^c$.Otherwise, set $\theta^j = \theta^{j-1}$.

5. Set $j = j + 1$ and return to step 2 until enough draws are obtained.

Once convergence is reached, all simulation values are from the target posterior distribution and a sufficient number will be drawn so that all areas of the posterior will be also explored.

3.6 Model selection and comparison

Model selection is to select the best model among several choices based on an evaluation of the performance of the models. A widely used statistic for comparing models in a Bayesian framework is the Deviance Information Criterion. The deviance information criterion (DIC) is a hierarchical modeling generalization of the AIC (Akaike information criterion) and BIC (Bayesian information criterion, also known as the Schwarz criterion). It is particularly useful in Bayesian model selection problems where the posterior distributions of the models have been obtained by Markov chain Monte Carlo (MCMC) simulation. Like AIC and BIC it is an asymptotic approximation as the sample size becomes large. It is only valid when the posterior distribution is approximately multivariate normal.

Define the deviance as $D(\theta) = -2\log(p(y \mid \theta)) + c$, where y are the data, θ are the unknown parameters of the model and $p(y \mid \theta)$ is the likelihood function. C is a constant that cancels out in all calculations that compare different models, and which therefore does not need to be known.

The expectations $\overline{D} = E[D(\theta)]$ is a measure of how well the model fits the data; the larger this is, the worse the fit.The effective number of parameters of the model is computed as $pD = \overline{D} - D(\theta)$,where $\overline{\theta}$ is the expectations of θ. The larger this is, the better it is for the model to fit the data.The deviance information criterion can be described as: $DIC = \overline{D} + pD$

The idea is that models with smaller DIC should be preferred than models with larger DIC. Models are penalized both by the value of $\overline{D}$,which favors a good fit, but also (in common with AIC and BIC) by the effective number of parameters pD.Since $\overline{D}$ will decrease as the number of parameters in a model increases, the pD term compensates for this effect by favoring models with a smaller number of parameters.

The advantage of DIC over other criteria, for Bayesian model selection, is that the DIC is easily calculated from the samples generated by a Markov chain Monte Carlo simulation. AIC and BIC require calculating the likelihood at its maximum over θ, which is not readily available from the MCMC simulation. But to calculate DIC, simply compute $\overline{D}$ as the average of $D(\theta)$ over a sample values of θ, and $D(\overline{\theta})$ as the value of D evaluated at the average of the samples of θ. Then the DIC follows directly from these approximations.

3.7 Tests for Convergence

It is important to establish whether a sequence of Markov chain Monte Carlo iterations has converged, that is, reached its stationary distribution. To examine the convergence of MCMC, considering a different method will be useful for detecting poorly sampled Markov Chains. Among several ways of a test of convergence, the most popular and straight forward convergence assessment methods will be used for this study. The following four methods are more likely to be considered for this study:

1. **Autocorrelation:** High correlation between the parameters of a chain tends to give slow convergence, where as high autocorrelation within a single parameter chain leads to slow mixing and possibly individual non convergence to the limiting

distribution because the chain will tend to explore less space in finite time. That is, low or high values indicate fast or slow convergence, respectively. In analyzing Markov chain autocorrelation, it is helpful to identify lags in the series in order to calculate the longer- run trends in correlation, and in particular whether they decrease with increasing lags.

2. **Time series plots or trace plots:** Time series plots (iteration number on x-axis and parameter value on y-axis) are commonly used to assess convergence. If the plot looks like a horizontal band, with no long upward or downward trends, then we have evidence that the chain has converged.

3. **Density plot:** The plots of all statistically significant covariates indicated that none of the coefficients have bimodal density and hence the simulated parameter values have converged.

4. **The Effective Sample Size:** The Effective Sample Size is a measure of efficiency that provides an estimate of the equivalent number of independent observations that are contained in the chain; this will of course be directly related to the degree of autocorrelation or dependence in the sequence for that parameter. A related concept to the MCMC convergence would be the inefficiency factor which is useful to measure the efficiency of the MCMC sampling algorithm. It is given as:- Inefficiency factor $= 1 + 2\Sigma_{k=1}^{\infty}\rho(k)$,where $\rho(k)$ is the sample autocorrelation at lag k calculated from sample draws. A large value of efficiency factor indicates that we need large MCMC iteration. The effective sample size, the number of MCMC output divided by the inefficiency factor. Let the output of MCMC denoted by N, then it can be given as: Effective sample size $= \frac{N}{1+2\Sigma_{k=1}^{\infty}\rho(k)}$.

Assessing Model Accuracy

After model convergence would be achieved, we need to run the simulation for a further number of iterations to obtain samples that can be used for posterior inference. The more samples we save, the more accurate would be our posterior estimates. One way of assess the accuracy of the posterior estimates is by calculating the Monte Carlo error for each parameter. This is an estimate of the difference between the mean of the sampled values (which we are using as our estimate of the posterior mean for each parameter) and the true posterior mean. As a rule of thumb, the simulation should be run until the Monte Carlo error for each parameter of interest is less than about 5% of the sample standard deviation.

4 Results and Discussion

4.1 Descriptive Result

In this study, the researcher answers the basic research questions and attained to address the objectives by modeling the data with the appropriate model fit. In order to study further the model, we were began from the simplest frequency table which has the power to intend the appropriate candidate model.

4.1.1 Summary of the data

This study was carried out to investigate the determinant factors associated with acute respiratory infections among under five children through analyzing the demographic, socio-economic, health, environmental and nutritional factors which were considered in similar studies conducted previously and using the geo-referenced data obtained from Ethiopian Demographic Health Survey of 2016. Accordingly, the study used 9,918 from 2016 EDHS and the two weeks prevalence of ARI among under-five children was about 8.4% in Ethiopia.

The major demographic and socioeconomic background characteristics of the respondents and children are presented in Table 4.1. Out of 9,918 under-five children, 51.04% and 48.96% were males and female respectively. The prevalence of ARI among males and females were 8.5% and 8.38% respectively. About 81.12% of under-five children were born to mothers who were residing in a rural area and had a higher prevalence of ARI (8.95%) as compared to under-five children who were born to mothers who lived in urban areas (6.25%).

Table 4.1 also showed that the majority of under-five children were born to the mothers in the age group 20-34 years with lower prevalence of ARI (8.55%) compared to the children were born to the mother in the age of 15-19 years (8.70%). The proportion of under-five children suffered from ARI varies with the household economic status. The highest incidence of ARI among under five children was found from a low income household, whereas the lowest proportion of under-five children suffered from ARI was recorded for those from higher income family.

Similarly, Table 4.1 shows that, among 9,918 respondents, (27.78%) of them had worked and the prevalence of ARI in under-five children whose mother had work was (9.33%). Additionally the proportion of children who had ARI varied by educational status of mothers. The highest proportion of the children with ARI was observed for children whose mothers had no education was (64.07%) with the prevalence of ARI were (9.60%) as opposed to the low prevalence of the ARI among under-five children, which was recorded for children whose mothers had secondary and higher educational level was (10.72%) with the prevalence of ARI (4.80%). Regarding with the age of under five children the prevalence of suffering from ARI were 12.3%, 11.39%, 8.43%, 8.29%, 7.44%, 4.85% for childrens whose age are 6-11,12-23, less than 6 ,24-35,36-47 and 48-59 respectively. This show as 6-11 month child and 12-23 month child had the highest percentage of ARI.

Also Table 4.1 shows that, the proportion of children who had ARI varied from one region to another. For example the highest prevalence of ARI was observed in Tigray (15.31%) followed by Oromia (14.40%) as opposed to the low prevalence which was recorded in Benishangul gumuz (2.58%) in two weeks preceding the survey date. Hence, there appears to be some variation in prevalence of ARI among the region of Ethiopia.

Table 4.1: Distribution of Demographic and Socio-economic Factors on prevalence of ARI among under-five children in Ethiopia (EDHS, 2016).

		ARI Status		
Variable Names	Category	Had no ARI	Had ARI	Total
Sex of child	Male	4632(91.5%)	430(8.5%)	5062(51.04%)
	Female	4449(91.62%)	407(8.38%)	4856(48.96%)
Place of residence	Rural	7325(91.05%)	720(8.95%)	8045(81.12%)
	Urban	1756(93.75%)	117(6.25%)	1873(18.88%)

Maternal age	15-19	336(91.30%)	32(8.70%)	368(3.71%)
	20-34	6565(91.45%)	614(8.55%)	7179(72.38%)
	35-49	2180(91.90%)	191(8.05%)	2371(23.90%)
Wealth index	Poorest	3325(90.50%)	3350(9.50%)	3675(37.05%)
	Poorer	1493(89.72%)	171(10.28%)	1664(16.78%)
	Middle	1272(92.17%)	108(7.83%)	1380(13.91%)
	Richer	1134(93.33%)	81(6.67%)	1215(12.25%)
	Richest	1857(93.60%)	127(6.40%)	1984(20.00%)
Mother occupation	Not working	6583(91.90%)	580(8.10%)	7163(72.22%)
	Working	2498(90.67%)	257(9.33%)	2755(27.78%)
Mother education	No education	5806(91.40%)	548(9.60%)	6354(64.07%)
	Primary	2261(90.40%)	240(8.60%)	2501(25.22%)
	Secondary and above	1012(95.20%)	249(4.8%)	1063(10.72%)
Child age	$<$ 6 months	1022(91.57%)	94(8.43%)	1116(11.25%)
	6-11 months	891(87.70%)	125(12.3%)	1016(10.24%)
	12-23 months	1704(88.61%)	219(11.39)%)	1923(19.39%)
	24-35 months	1748(91.71%)	158(8.29%)	1906(19.22%)
	36-47 months	1755(92.56%)	141(7.44%)	1896(19.22%)
	48-59 months	1961(95.14%)	100(4.85%)	2061(20.78%)

Number of child	1-3 child	4551(92.00%)	397(8.00%)	4948(49.89%)
	4-6 child	3226(91.05%)	317(8.95%)	3543(35.72%)
	above 6 child	1304(91.40%)	123(8.60%)	1427(14.39%)
Region	Tigray	835(84.69%)	151(15.31%)	986(9.94%)
	Affar	892(92.44%)	73(7.56%)	965(9.73%)
	Amhara	824(89.18%)	100(10.82%)	924(9.32%)
	Oromia	1272(85.60%)	214(14.40%)	1486(14.98%)
	Somale	1329(95.68%)	60(4.32%)	1389(14.00%)
	Benishangul gumuz	793(97.42)%)	21(2.58%)	814(8.21%)
	SNNPS	1081(90.31%)	116(9.69%)	1197(12.07%)
	Gambela	624(95.41%)	30(4.59%)	654(6.59%)
	Harari	543(97.31%)	15(2.69%)	558(5.63%)
	Addis Ababa	411(94.05%)	26(5.95%)	437(4.41%)
	Dire Dawa	477(93.90%)	31(6.10%)	508(5.12%)

The major health, environmental and nutritional related background characteristics of the respondents and children are presented in Table 4.2. It shows that there is the highest prevalence of ARI was recorded for those households used unsafe fuel for cooking (8.60%) compared to the lowest proportion of ARI recorded for those households used safe fuel for cooking (5.30%). Among 9,918 under-five children, 43.10% of them received vitamin A recently with the lowest proportion on the prevalence of ARI (7.75%) compared to not having vitamin A recently (9%). Additionally, 11.13% of under-five children had Diarrhea recently with the highest prevalence of ARI (24.64%)

compared to not having Diarrhea recently (6.41%).

The proportion of under five children suffered from ARI also differ from the source of water in their household used. The highest prevalence of ARI was observed for the child whose source of drinking water were unprotected/unimproved (9.39%).From the nutritional status of children like stunting and wasting, the result indicates that 36.10% of under-five children were stunted (chronic malnutrition) and 11.46% of under-five children were wasted (acute malnutrition). The prevalence of ARI in under-five children with stunting and wasting were 10.59%, 10.03% respectively.

Table 4.2: Distribution of Health, Environmental and Nutritional related Factors on prevalence of ARI among under-five children in Ethiopia (EDHS, 2016).

		ARI Status		
Variable Names	Category	Had no ARI	Had ARI	Total
Vitamin A supplement	No	5134(91.00%)	505(9.00%)	5639(56.90%)
	Yes	3947(92.25%)	332(7.75%)	4279(43.10%)
Breast feeding	Never breast feed	358(94.46%)	21(5.54%)	379(3.82%)
	Not currently	4932(93.25%)	357(6.75%)	5289(53.33%)
	Still breast feed	3791(89.20%)	459(10.80%)	4250(42.85%)
History of diarrhea	No	8249(93.59%)	565(6.41%)	8814(88.87%)
	Yes	832(75.36%)	272(24.64%)	1104(11.13%)
Type of cooking fuel	Unclean/unsafe	8599(91.40%)	810(8.6%)	9409(94.87%)
	Clean/Safe	482(94.70%)	27(5.30%)	509(5.13%)
Wasting	Not wasted	8058(91.80%)	723(8.20%)	8781(88.50%)
	Wasted	1023(89.97%)	114(10.03%)	1137(11.50%)

Stunting	Not stunted	5880(92.77%)	458(7.23%)	6338(63.90%)
	Stunted	3201(89.41%)	379(10.59%)	3580(36.10%)
Source of water	Unprotected	5182(90.61%)	537(9.39%)	5719(57.66%)
	Protected	3899(92.85%)	300(7.15%)	4199(42.34%)

The bivariate association between ARI status of under-five children and predictors were showed in Appendix A:Table 5.1 indicates that ARI status was strongly associated with region, age of child, wealth index of household, place of residence, breastfeeding, maternal educations, vitamin A supplement, had diarrhea recently, type of cooking fuel, maternal work, wasting, stunting and source of drinking water used were found to have a significant association with the two weeks incidence of ARI at the 5% significance level.

4.2 Test of Heterogeneity.

Before analyzing multilevel data, one has to test the heterogeneity of ARI status among regional states of Ethiopia from which essential clues would be obtained for incorporating the random effects. Therefore, the Pearson chi-square for the proportion of acute respiratory infections among under five children across the region has been investigated in the Table 4.3. Consequently, as it obtained by cross tabulations the Pearson Chi-square (χ^2) = 249.058 with 10 degrees of freedom. The P-value is less than 0.05 level of significance, implying strong evidence of heterogeneity for the ARI status of under-five children across regional states of Ethiopia. Hence we have enough evidence to reject the null hypothesis and conclude that there is heterogeneity of acute respiratory infections among under five children in regional states of Ethiopia.

Table 4.3: Heterogeneity of ARI of under-five children between Regional States of Ethiopia.

Chi-square test			
Statistics	Value	D.f	P-value
Pearson Chi-square	249.058	10	0.000*
N of valid cases 9918			

4.3 Bayesian Multilevel Logistic Regression model

Bayesian multilevel logistic analysis procedure was used to make inference about the parameters of a multilevel logistic model. Bayesian procedure was considered in this study to make inference about the parameters of a logistic regression model. The Bayesian method gives estimates of parameters by sampling them from their posterior distributions through an MCMC method. This approach was employed to model ARI status of under-five children. The metropolis Hasting algorithm were implemented using non-informative uniform prior distribution with scale parameter (0, 1) for the fixed effects and inverse gamma distribution with a scale of 0.001 and shape 0.001 for random effect (Browne, 2009). The data used in this study have a hierarchical structure. Here, the level-1 units are the individual under-five children, and the level-2 units are the regions that constitute the groups into which the children are clustered or nested.

In Bayesian multilevel models,three models were fitted. Model 1 does not include any covariate; Model 2 includes fixed effects and regional random effects; and Model 3 includes variables in Model 2 and variable that has a significant impact on ARI among under five children in the Bayesian multilevel intercept model by observing their respective region effect to identify the appropriate model which fits our data.

4.4 Model comparison of Bayesian multilevel logistic regressions.

From Table 4.4 we see that the comparison of the fit of Bayesian multilevel logistic regression models using the summary of the fitted model. The researcher is going to compare Bayesian multilevel empty model, Bayesian multilevel random intercept model and Bayesian multilevel random coefficient model using DIC further strengthened the advantage of the Bayesian multilevel model. The model which has small DIC is the best model for the data set of under-five child ARIs in Ethiopia.

The result, shows that Bayesian multilevel random coefficient model was an improved fit as compared to the rest models in any combination of variables in the data set. The DIC diagnostics of random intercept Bayesian multilevel logistic regression model are reduced by 375.16 from the Bayesian multilevel logistic regression of an empty model. This show as adding covariate variables to the model indicates how the variable was determined the occurrence of ARIs. Thus; Bayesian multilevel logistic regression for random intercept was the better model as compared to Bayesian multilevel for an empty model. DIC is a composite measure of how well the model does (it is a compromise between fit and complexity), and small values of DIC is preferred. As we can see in Table 4.4 the Bayesian multilevel random coefficient model has got the smallest DIC value so, its the best model to determine the incidence of ARI among under-five children in Ethiopia.

Therefore, this Bayesian deviance information criterion showed that Bayesian multilevel random coefficient model is the more significant model and best fit the data. The average deviance from the complete set of iterations ($\overline{D}$) also decreased from an empty model to random intercept and from random intercept to the random coefficient model. $D(\overline{\theta})$ shows that the deviance at the expected value of the unknown parameters and it also shows the decreasing trend from an empty model to random intercept and from random intercept to the random coefficient model. Also the model complexity is measured by pD (The effective number of parameters in the model), the larger the pD is easier to fit the data. Based on this fact, the third model has the largest value of this measure, it is selected again.

Table 4.4: DIC values for model comparisons.

BDIC for model comparison				
Model	$\overline{D}$	$D(\overline{\theta})$	pD	DIC
BML logistic regression of Empty model	5097.64	4848.33	249.31	5346.95
BML logistic regression of Random intercept model	4699.42	4427.04	272.38	4971.79
BML logistic regression of Random coefficient model	4612.55	4275.15	337.40	4949.95

Note: BDIC: is Bayesian Deviance Information Criterion

BML: is Bayesian Multilevel

$\overline{D}$: is the posterior mean of the deviance, measuring how well a model fits the data.

$(\hat{\theta})$: The deviance at the expected value of the unknown parameters.

pD: is the effective number of parameters measuring model complexity.

4.5 Result of Bayesian multilevel empty model

We first fit a simple model with no covariates i.e. an intercept-only model that predicts the probability of the ARI status of under five children. The simplest 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 contains no explanatory variables and it can be considered as a parametric version of assessing heterogeneity among regions with respect to the ARI status of under five children. From Table 4.5 below, both data showed that there is a significant variation among the region. The regional variation of the ARI status of under five children was significant.

Table 4.5: Posterior summaries for parameters of the empty model.

Model	Post. Mean	MC error	SD	2.5%	50%	97.5%
Fixed intercept(β_{0j})	-2.789	0.0003	0.072	-2.936	-2.789	-2.652
Random intercept var(U_{0j})	0.998	0.0006	0.136	0.750	0.989	1.279

From Table 4.5, the overall posterior mean of ARI status of under five children without incorporating the covariate is estimated to be $\beta_{0j} = -2.789$ and the between-region (level two) variance of ARI status of under five children is estimated as $\sigma^2_{uo} = 0.998$ which is found to be significant because the credible interval of the respective parameters was greater than zero. Indicating the variations of ARI status of under five children among regional states of Ethiopia.

Here the null hypothesis tested is $\sigma^2_{0u} = 0$. i.e., there is no regional variation in the incidence of ARI among under five children in Ethiopia. Based on the above result, the values are significant at 95% credible interval, which means that the interval is greater than zero, therefore, the null hypothesis had been rejected indicating strong evidence that the between region variance is greater than zero. The variance of the random factor is significant which indicates that there are regional differences in prevalence of ARIs among under five children.

The variance of σ^2_ϵ and σ^2_{0u} estimate the variations among individual under five children and among regions of the country respectively. In the variance component model it is possible to decompose the variance in to regional level (higher level) and individual level. Individual (level-1) variance was to assess how much of the variation is due to the individual themselves and how much of the variation is due to regional level.

In order to get an idea of how much of the variation in ARI status of among under five children was attributable to the region level factors, it is useful to see the intra-region correlation coefficient (ICC) as $\hat{\rho}=\frac{0.998}{0.998+3.29} = 0.233$,which measures the proportion of variance of the ARI of under five children that is between regions, not within regions. This means that around 23.3% of the variance in ARI of under five children are due to variation between regions. Whereas the remaining 76.7% attributable to individual level, i.e., within regional differences.

4.6 Results of Random Intercept Bayesian multilevel model

Results of analysis of the Bayesian multilevel random intercept model are displayed in (Appendix B; Table: 5.2). Intercept and some covariates are significant. In this Bayesian intercept model, the intercept is allowed to vary across the region after incorporating covariates of the ARI status of under-five children. This means that, the intercept β_0 is shared by all regions, while the random effect U_{0j} is specific to region j and the random effect is assumed to be a normal distribution with variance σ_u^2.

The overall posterior mean of ARI status is estimated to be -3.280 which decreased by 0.491 as compared to an empty model (Table: 4.5).Therefore, indicating many variables that are included in this model have impacts on the incidence of ARIs.

The result shows that the variance of the random effect is significant which indicates that there are regional differences in ARI status of under five children in the given data set. The Bayesian multilevel logistic regression analysis result of intercept model displayed in (Appendix B; Table 5.2) also estimates the variance of random effect at the regional level $var(U_{0j}) = 1.107$ since the 95% credible interval was greater than zero under the interval, which indicates that there is a significant regional variation. This confirmed the significance of the regional difference in prevalence of ARI status among under five children in Ethiopia.

In general, the variance component for random intercept is found to be significant because the lower limit of the credible interval is greater than zero, indicating strong evidence of the variations across regions for prevalence of ARI.

The results displayed in (Appendix B; Table: 5.2) showed that the intra-region correlations coefficient (ICC) is estimated as $\hat{\rho}=\frac{1.107}{1.107+3.29} = 0.252$. This means that about 25.2% of the total variability in ARI status of under-five children is due to difference across regions, with the remaining unexplained 74.8% attributable to individual differences.

4.7 Results of Random Coefficient Bayesian Multilevel model

Results of analysis of the Bayesian random coefficients model are displayed in Table 4.6. Bayesian multilevel random coefficient logistic regression analysis of ARI status was compared with an empty and intercept model based on their respective Deviance information criteria. In this context the deviance information criteria for the random coefficient model was smaller. This suggests that the model with all predictors, including the random slope model is found to be better than the intercept models because the DIC for the random coefficient model was smaller. Generally, the Bayesian multilevel random coefficient was the lowest, showed that the random coefficients model was appropriately fitted the data.

The results revealed that, age of child, wealth index of household, maternal education level, vitamin A supplement, had diarrhea recently, maternal working status, stunting and source of drinking water, were found to be significant, indicating strong effects on ARIs of under five children and also contributing to variations among regional state in Ethiopia. However, the impacts of sex of the child, maternal age, place of residence, breast feeding status, type of cooking fuel, wasting and number of living child are found to be insignificant, according to this study no evidence for the effects of those factors on ARIs of under-five children. From the output of the random coefficient Bayesian multilevel models, we interpret the results as follows:

The child age has a significant association with incidence of ARIs. The odds of incidence of ARI among under five children in the age group of 48-59 month, were 40.2% (OR=0.598, 95% Credible interval:(-0.908,-0.130)) times less likely having ARI to those children in the age group less than 6 months. This means that as the age of child increase the ability of children being suffered from ARI was decreased.

According to this study, wealth index of household also showed a statistical significant association with incidence of ARI. The odds of the child having ARI who comes from middle class family was 46.8% (OR=0.532, credible interval:(-0.927, -0.343)) less likely than the child comes from poorest family. The odds the child having ARI who comes from a richer class family was 58.4% (OR=0.416, credible interval :(-1.417, -0.435)) less likely than the child comes from poorest family. Similarly, the odds the child having ARI who comes from richest class family was 47.2% (OR=0.528, credible interval :(-1.051, -0.226)) less likely than the reference group. From this, we conclude that the child comes from low income families was the most significantly being suffered from ARI in Ethiopia. From this study researcher demonstrated that children born in lower socioeconomic groups suffered more from ARI than the children born in higher socioeconomic groups. It indicates that family income had a significant association with the prevalence of ARIs.

The odds of maternal education level have a significant association with ARIs. The odds of under-five children developing ARI among mothers who completed secondary and above educational levels to have ARI was 33.5% (OR=0.635, credible interval: (-0.875, -0.067)) times less likely than under five children affected by ARIs as compared with children whose mothers had no education.. This means that a child whose families are not educated were mostly affected by ARIs.

According to this study Vitamin A consumption of children in the last six months was found that, it had a significant effect in the ARIs. The odds of a child who had received vitamin A recently was 16.6% (OR=0.834, credible interval: (-0.359, -0.011)) less likely to suffer in ARI compared to a child not received vitamin A recently. Indeed, the odds of Under-five child who had Diarrhea recently were 4.284 (OR=4.284, credible interval :(1.260, 1.649)) times more likely a child who have ARI than who had not Diarrhea recently.

Maternal occupation also has a statistically significant association with ARI; accordingly, the odds of children whose mothers had occupation, to have ARI incidence was increased by 30.9%(OR: 1.309, CI (0.011, 0.540)) compared to mothers had no occupations. Another finding from the Table 4.6 showed that, chronic nutrition status of children explained by stunting was included in the model and it has a significant effect on ARI. The odds of under-five children who were stunted were 59.5% (OR=1.595, credible interval: (0.291, 0.640)) times more likely to experience ARI than who were not stunted.

This study also revealed that, the source of drinking water has a significant effect on ARI. The odds of children using protected source of drinking water were 18.6% (OR=0.814, credible interval :(-0.40, -0.016)) less likely than having ARI as compared with a child using unprotected source of drinking water.

The Bayesian multilevel logistic regression analysis of random coefficient model results displayed in Table 4.6 below, also estimates the variance of random effect at the regional level, $var(U_{0j})$. Thus, the value of var $(U_{0j}) = 1.295$ indicate there was significant variation (which means the 95% credible intervals is greater than zero). This confirmed the significance of the regional difference in incidence of ARIs in the regional state of Ethiopia. The researcher tried to identify to see the level of variation; that the intra-region correlation coefficient ICC is estimated as $\hat{\rho} = \frac{1.295}{1.295+3.29} = 0.282$. This means that about 28.2% of the total variability in incidence of ARI are due to differences across regions, with the remaining unexplained 71.8% attributable to individual differences.

In the random intercept model, we allowed the intercept only to vary across regions by fixing explanatory covariates. In this model, researcher has tested the variable that have significant impact on occurrence of ARI among under five children in the Bayesian multilevel intercept model by observing their respective region effect. Consequently, regional level variables which are supposed to varying regionally such as wealth index of household and maternal working status have been examined.

The region wise intercept (U_{0j}), wealth index slopes (U_{u11}) and maternal work slopes (U_{u21}) vary significantly. There was a significant variation in the effects of these explanatory variables across the regions. The researcher revealed that the variance of the wealth index of richer category has slopes $(\sigma^2_{u11} = 0.725)$ with a credible interval of (95% CI: 0.223, 1.606) and the variance of the mothers who had worked slopes $(\sigma^2_{u21} = 0.502)$ with (95% CI: 0.188, 1.031) the interval was greater than zero. This indicates that the random slopes of wealth index of household and the maternal working status across the region was significant. This means that the wealth index of household and maternal working status for incidence of ARI among under five children varies from region to region.

Table 4.6: Bayesian Estimates for Random coefficient model.

Fixed effects								
Variables	category	Coefficient	MC error	SD	Exp(β)	2.5%	50%	97.5%
–	Intercept	-3.305	0.0057	0.353	0.0367	-4.025	-3.307	-2.615
Child age	<6 (ref)							
	6-11	0.259	0.0008	0.163	1.295	-0.061	0.263	0.579
	12-23	0.095	0.0008	0.151	1.099	-0.199	0.094	0.388
	24-35	-0.155	0.001	0.175	0.856	-0.489	-0.157	0.188
	36-47	-0.164	0.0011	0.190	0.849	-0.541	-0.159	0.192
	48-59	-0.514	0.0011	0.201	0.598	-0.908	-0.517	-0.130
Wealth index	Poorest (ref)							
	Poorer	-0.222	0.0005	0.125	0.800	-0.472	-0.224	0.021
	Middle	-0.632	0.0006	0.147	0.532	-0.927	-0.629	-0.343
	Richer	-0.877	0.0015	0.245	0.416	-1.417	-0.863	-0.435
	Richest	-0.638	0.001	0.210	0.528	-1.051	-0.635	-0.226
Maternal educ.	No education(ref)							
	Primary	0.140	0.0004	0.106	1.150	-0.064	0.140	0.350
	Secondary and higher	-0.454	0.0007	0.205	0.635	-0.875	-0.459	-0.067
Vitamin A	No (ref)							
	Yes	-0.182	0.0003	0.088	0.834	-0.359	-0.183	-0.011
Had diarrhea	No(ref)							
	Yes	1.455	0.0003	0.099	4.284	1.260	1.455	1.649
Maternal work	Not working(ref)							
	Working	0.270	0.0007	0.137	1.309	0.011	0.271	0.540
Stunting	Not stunted(ref)							
	Stunted	0.467	0.0003	0.089	1.595	0.291	0.464	0.640
Source of water	Unprotected(ref)							
	Protected	-0.206	0.0004	0.098	0.814	-0.400	-0.204	-0.016
	Random intercept							
	σ^2_{u0}	1.295	0.0012	0.203	-	0.925	1.284	1.724
	Random slope							
	σ^2_{u11}	0.725	0.0049	0.351	-	0.223	0.666	1.606
	σ^2_{u21}	0.502	0.0032	0.223	-	0.188	0.468	1.031

Note: Ref - is reference category

σ^2_{u0}=regions variance,σ^2_{u11}=wealth index variance, σ^2_{u21}=maternal working variance

4.8 Assessment of model Convergence

Once a model has been developed, we now would like to know how effective the model is in describing the outcome. This is referred to as goodness of fit. These are a method used to determine whether the algorithm has reached its equilibrium or target distribution. There are several ways used to monitor convergence. The most common ways of checking goodness of fit are: diagnosis for convergence of the MCMC chains was confirmed by visual inspection the trace plot, kernel density, effective sample size and Monte Carlo (MC) error. Small values of the MC error indicate that the quantity of interest has been calculated with precision.

To use summary statistics of the estimated posterior distributions for inference the realized value of the parameters (the MCMC value) should converge. To check this we have to use a suitable diagnosis to evaluate mixing and convergence of a sampler. From different methods of checking convergence, trace and history plots, kernel density plot and autocorrelation are among the common (Ntzoufras, 2009). Tests used for checking convergence of a Bayesian multilevel random coefficient model were as follows:

Time Series Plots: are commonly used to assess convergence of the parameter estimates in Bayesian analysis. The plot with number of iterations on the x-axis and parameter values on the y-axis for each significant parameter. The plot looks like a horizontal band, with no longer upward or downward trends, then we have evidence that the chain has converged. For all simulated parameters, time series plot indicates a good convergence since the chains are mixed together (see below figure and Appendix D).

Kernel Density Plot: This is also the statistical techniques to recognize convergence in Bayesian analysis. The plot on below figure and Appendix D shows that the coefficients of significant variables were approximately normally distributed. Thus, this indicates that the Markov chain has attained its posterior distribution.

Autocorrelation Plot: It is a test used for convergence of Bayesian analysis. The ACF measures how correlated the values in the chain are with their close neighbors. The lag is the distance between the two chains to be compared. High autocorrelation indicates slow mixing within a chain and usually slow convergence to the posterior distribution. So, the plots displayed below indicate low autocorrelation as we have seen from below figure and Appendix D.

Effective sample size: convergence of posterior estimate has been checked using an effective sample size that is all the effective sample sizes of the estimates are greater than 200. So, the more samples you save, the more accurate would be your posterior estimates.This is an indication of efficient posterior estimate. It has been presented in below figure and appendix D.

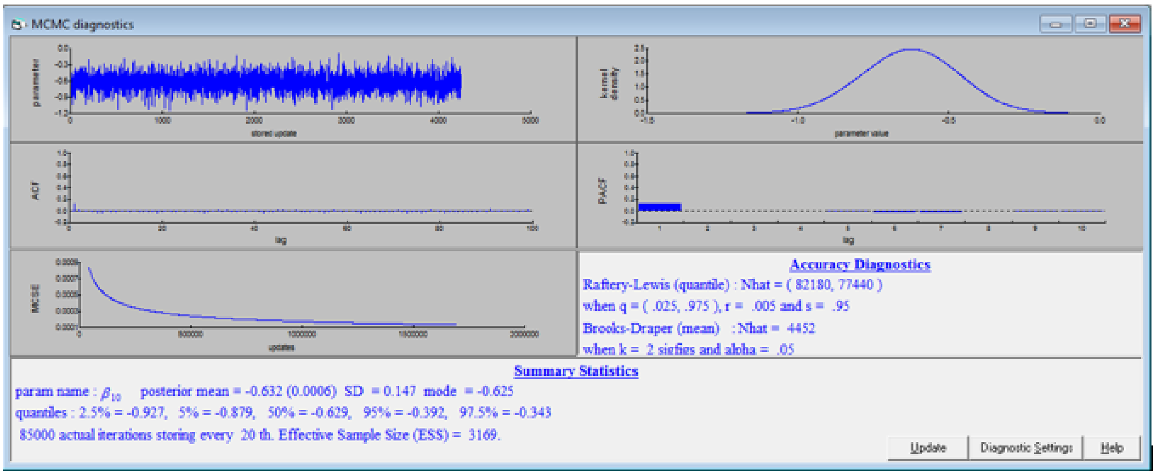


Figure 4.1: Convergence for wealth index of household in middle class

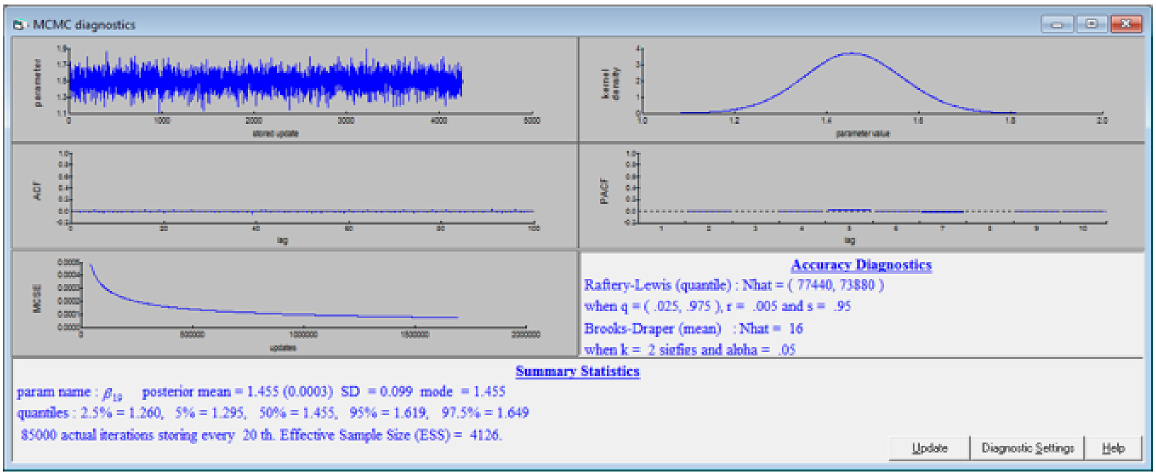


Figure 4.2: Convergence of a child had diarrhea recently

4.8.1 Assessing Accuracy of the Bayesian model Estimate

The posterior summary estimates by the MCMC algorithm (metropolis Hasting algorithms), like posterior mean, standard deviation, Monte Carlo error and credible interval were estimated using **MLwiN software**. To assess the accuracy of the Bayesian multilevel analysis, the researchers were used Monte Carlo error for each parameter. If the MC error value is less than 5% of its posterior standard deviation, then the posterior density estimates with accuracy. In this study, MC error for each significant variable is less than 5% times its standard deviation. This indicates that convergence and accuracy of posterior estimates are attained and the model is appropriate to estimate posterior statistics. Therefore, since the convergence and accuracy criteria were attained, then it is possible to say that the posterior estimate of the Bayesian random coefficient model was accurate

4.9 Discussions of The Results

The objective of this study was to identify factors associated with the prevalence of ARI among under-five children and to identify regional variation based on the 2016 EDHS using Bayesian multilevel logistic regression model. For determining the risk factors of ARI morbidity among under-five children; a total of 9,918 children were included in the study out of which 8.4% were suffered from ARI. This is consistent with the finding reported in the 2011 EDHS. According to the descriptive data analysis, the highest prevalence of ARI was observed in Tigray (15.31%) followed by Oromia (14.40%) as opposed to the low prevalence which was recorded in Benishangul regional state (2.58%). This could be due to several reasons for the regional variations in the disease distributions including environmental, climatic, and infrastructure difference across the regions. In addition, the difference in the altitude of the regions in Ethiopia could contribute to the difference in ARI.

For the analysis of the data, Bayesian multilevel logistic regression model was used. Firstly, the researcher has fitted the Bayesian multilevel empty model, intercept and coefficient models. In this regard, the Bayesian random coefficient model was found better in fitting the data appropriately based upon their deviance information criteria compared to the empty and random intercept model. In the multilevel analysis under five childrens are nested within various region in Ethiopia. Before the analysis of data using multilevel approach, the necessity of multilevel analysis was investigated through Chi-square test statistic. The heterogeneity test and the significance of variance of random effect suggest that ARI among under five children differs among regions.

The results which are obtained from the Bayesian multilevel logistic regression of random coefficient models are discussed as follows:The above analysis revealed that, age group of children between $48-59$ months has lower odds of ARIs as we compare with children less than 6 months, this means that as age of child increase the occurrence of ARI was reduced. The result was confirmed by similar findings investigated previously by Geberetsadik *et al.,* (2015) and studies done in Ahmadabad City. This is due to the fact that as children get older, their immunity grows stronger and becomes better able to resist infection.Also the result of this finding seems to agree with the previous study using systematic analysis has examined ALRI incidence was highest in neonates aged $0-27$ days and infants aged $0-11$ months as compared to the older age (Nair H, 2013). This finding was consistent with the results of the study in Wondo-Genet district, southern Ethiopia (Abuka T., 2016) and this finding was also consistent with the results of the study done in Rwanda (Harerimana *et al.,* 2016).

In the same manner mothers educational status was strongly associated with ARIs. Results of Bayesian multilevel logistic regression of random coefficient model show that mothers that are more educated become less fatalistic about their children's illnesses, they are more capable of seeking available health facilities and their greatly change the traditional balance of family relationships with profound effects on child care. Also, it revealed that caregivers with high education level were more likely to seek appropriate care for their children this finding is consistent with (Rehman, 2014) and (Noordam *et al.,* 2015). In addition, the result indicates that women whose education levels were secondary, and higher tend to have less probability of ARI among under-five children than those who have no education. Moreover, decreasing pattern in ARI has been observed when their education levels increases this is due to mothers with higher levels of education were more likely to seek immediate health facility care compared to those with lower levels of education. This finding was in line with Jabessa S (2015) and

Keter, (2015).

According to this study, the maternal work have also the significant determinant of the incidence of ARI among under-five children. Consequently, child whose mothers have work were more likely to suffer from ARIs than a child whose mothers have not worked. This logic might be occurred due to the fact that time allocated to earning income may be at expense, time spent in feeding and caring for children (When mothers are at work, they dont have sufficient time for breastfeeding their children; thus children become vulnerable to ARI). Moreover, since the majority of mothers in developing countries like Ethiopia works in the informal sector and in lower status jobs, the amount of income of these mothers is low and would have a significant impact on ARI incidence of children. The other reason could be that working mothers have been exposed to certain chemicals, pollutants, or toxic fumes in the working environment, thereby transmitting the infection to their children may be increased. The result of this finding seems to agree with the previous study of a similar case Jabessa S., (2015) and Fatmi Z, (2002).

Household wealth index also showed a statistical significant association with incidence of ARI. Being a higher wealth quintile, compared to low (poorest), reduce the probability of ARI occurrence. This finding is in line with studies done in Uganda Being a higher wealth quintile, compared to the lowest (poorest), reduce the probability of ARI occurrence by 5%-18% (Bbaale, 2011). This may due to families with high socioeconomic status are supposed to drink more piped water, and use hygienic toilets. Additional this finding revealed that, children from households with higher socioeconomic status were more likely to be taken to a health facility or to be taken to a hospital than those of low socioeconomic status. This result is consistent with previous studies done by Burton *et al.,* (2011), Keter, (2015) and Bhat, (2013).

The findings of this study also show that under-five child who had Diarrhea recently were 28.4% times more likely have ARI than under-five child who had not Diarrhea recently. This might be Diarrhea morbidity affect the child immunity of under-five children. This result also in agreement with (Dadi *et al.,* 2014) and also supported by a study done in Nepali, this was due to occurrence of ALRI increased as the number of days with diarrhea in the prior 28 days increased; the greatest incident rate was reported among children with 20 or more days of diarrhea in the preceding 28 days (Walker *et.al.,* 2013).

Under-five children who were chronic malnutrition (stunted) were 59.5% more likely to experience ARI than under-five children who were not stunted. This may due to wasting in turn impairs the function of the immune system and can lead to increase the severity, duration of and susceptibility to acute respiratory infection or due to lower intake of adequate food, improper treatment and insufficient care giving a child may exposed to ARIs. This finding is in line with Ramadhani, (2013), Geberetsadik *et al.,* (2015) and Bbaale, (2011).

A child who had received vitamin A recently was 16.6% less likely to suffer in ARI compared to a child not received vitamin A recently. In fact to reduce the incidence of ARI among under five children and to reduce the childhood death due to ARI, health workers promote Vitamin A supplementation for all in this age group. Family members have to seriously supervise their childs Vitamin A intake. This finding was in line with (Azad, 2009). But this finding was inconsistent with Geberetsadik *et al.,* (2015) findings.

Another finding was showed, that source of drinking water was an important determinant of the incidence of ARIs. The odds of children using protected source of drinking water were 18.6% less likely than having ARI as compared with a child using unprotected source of drinking water. Therefore, the results indicate that source of drinking water is associated with the ARI status of a child through its impact on the risk of childhood diseases such as diarrhea, and are affected indirectly as a measure of wealth and availability of water. Use of improved water and sanitation has a lot of benefits in the reduction of diseases (particularly ARI). This result was supported by similar findings by (Briggs, 2003; Ghimire M, 2010).

The analysis based on Bayesian multilevel logistic regression provided estimates for variances of the random effects and interclass correlations. The estimates for each level were different, suggesting that the variance composition of ARI status was different at individual and regional levels. This means that the sources of variations are individuals and regions. The result of Bayesian multilevel logistic regression model comparison indicates that, the random coefficient Bayesian multilevel logistic regression model best fits the data than the null model and intercept model of Bayesian multilevel logistic regression model and the interpretation was depend on random coefficients.

The model would be implemented using **MLwiN 2.02 versions** (Rasbash et al., 2005). For each model, 85,000 MCMC iterations were run, with the initial 15,000 burn-in terms discarded, and thereafter keeping every 20th sample value to make observations independent or low autocorrelations. The 60,000 iterations left were used to assess convergence of the chain and parameter estimation. For convergence check researcher has implemented four methods such as trace plot, density plot, autocorrelation, and effective sample size. With these four methods the convergence of the posterior estimate was correctly achieved and MC error for each significant predictor was found to be less than 5% of its posterior standard error.

5 Conclusions and Recommendations

5.1 Conclusions

The purpose of this study was to investigate factors associated with prevalence of ARI among under five children and to identify regional variation based on data from 2016 EDHS. This study found evidence that verifies some demographic, socioeconomic, health, nutritional and environmental variables considered in this study have significant influence on the ARI morbidity among under-five children. Factors such as: Age of child, wealth index of household, womens educational level, vitamin A supplement, mothers occupations, had diarrhea recently, chronic malnutritions and source of drinking water were found to be determinant of incidence of ARI among under five children.

This study demonstrated that children born in lower socioeconomic groups suffered more from ARI than the children born in higher socioeconomic groups and a child who had diarrhea was more suffered from ARI. This study also revealed that prevalence of ARI among under-five children was less likely for children whose mothers are from secondary and above educational level as compared to those of the uneducated and the risk of ARI decreases as the age of the child increases. The predictors, like vitamin A supplement is another, significantly associated with the incidence of ARI among under five children and as this analysis, the probability of child suffered from ARI was less for children had vitamin A as compared to those who do not receive Vitamin A. Based on source of drinking water, the odds of children using an improved source of drinking water were less likely to suffer from ARI compared to those who do not use improved source. Likewise, the odds of children having mothers had occupation, to have ARI symptoms was increased as compared to mothers having no occupation and stunted children also more likely to have ARI.

An inference is a fully Bayesian multilevel model based on recent Markov chain Monte Carlo techniques. Based on DIC, Bayesian multilevel random coefficient model was appropriately fitted the data than empty and intercept model. The effect of regional variations in the wealth index of household and maternal working status further implies that there exist considerable difference in prevalence of ARI among regions and a model with a random coefficient is more appropriate to explain the regional variation than a model with empty and intercept model.

Finally we conclude that, there is a huge regional disparity in the magnitude of ARI in children across the regions in Ethiopia. This may suggest differences in lifestyle, culture, ethnic or environmental determinants between different regions, because of these potential cultural, socioeconomic, and environmental differences of ARI among under five children's exhibits a significant variation among regional states of Ethiopia.

5.2 Recommendations

Based on finding and results of this study, the following recommendations can be made:

- Educational level of mothers plays an important role to save children from ARIs symptoms. This is, however, requires a long-term investment. As an alternative, in the short term, health programs need to focus on supporting women with little or no education.

- The government should design programs to improve the socioeconomic standards of the poor, there is a clear need for intervention to reduce economic inequalities and ultimately poverty among the populace. Higher incomes mean that households can afford improved water source and better cooking fuels.

- The government needs to implement an integrated package of protecting children against diarrhea and lack of Vitamin A supply and other child care programs to reduce ARI among children.

- Among under-five children incidence of ARI symptoms between regions are significant. This is an indication that the severity of ARI varies from one region to another. Thus, in order to have a bearing on policy recommendations, future studies should focus on identifying risk factors of incidence of ARI symptoms for each region of Ethiopia separately.

- Further studies may conducted to identify other factors that affect respiratory systems of under five children among regions.

- Although the variation across the regions has been addressed with this study, the distribution for the prevalence of ARI and the issue of identifying the hotspot-area is not covered here. Therefore, the researchers are recommended to extend this study with the application of spatial models.

- Finally; in this study researcher used the MCMC estimation methods for simulation which is based on the iteration technique. However, there is also another recent, deterministic and very fast estimation method called Integrated Nested Laplace Approximation (INLA) method. Hence, researchers should strongly recommend so that to compare the MCMC and INLA for better estimation of the posterior marginal.

5.3 Limitations of the Study

The data used here being secondary may have a number of limitations on the outcome of this study. The following limitations have to be kept in mind:

- Information was based on mothers reporting about the ARI status of children and no clinical examination were undertaken, which may have caused misclassification bias. Since the study also was cross-sectional in design, it may not be strong enough to demonstrate direct cause-and effect relationships between risk factors and ARI.

- The study focused on identifying the factors that associated with the prevalence of ARI among under five children in Ethiopia. However, couldnt incorporate some of the most influential factors such as climate, seasonality and house ventilation were not found in the data set.

References

Abuka T. (2016). Prevalence of pneumonia and factors associated among children 2-59 months old in Wondo Genet district, Sidama zone, SNNPR, Ethiopia. Current Pediatric Research.

Acquah, H. D. (2013).Bayesian logistic regression modelling via markov chain Monte Carlo algorithm.*Journal of Social and Development Sciences,* 4(4), 193-197.

Azad, K. M. A. K. (2009).Risk factors for acute respiratory infections (ARI) among under-five children in Bangladesh.*Journal of Scientific Research, 1*(1), 72-81.

Bernardo, J. M., Smith, A. F. (2009).*Bayesian theory (Vol. 405).* John Wiley Sons.

Bhat, R. Y., Manjunath, N. (2013). Correlates of acute lower respiratory tract infections in children under 5 years of age in India.*The international journal of tuberculosis and lung disease, 17*(3), 418-422.

Bornmann, L., Mutz, R., Daniel, H. D. (2013).Multilevelstatistical reformulation of citationbased university rankings:The Leiden ranking 2011/2012.*Journal of the American Society for Information Science and Technology, 64*(8), 1649-1658.

Briggs, D. (2003).*Making a difference: Indicators to improve children's environmental health*(pp. 1-17).

Browne, W. J. (2017). MCMC estimation in MLwiN.

Browne, W. J., Rasbash, J. (2009).*MCMC estimation in MLwiN.* Bristol: Centre of Multilevel Modelling, University of Bristol.

Browne, W. J. (1998). Applying MCMC methods to multilevel models. PhD thesis.*University of Bath, UK.*

Burton, D. C., Flannery, B., Onyango, B., Larson, C., Alaii, J., Zhang, X., Feikin, D. R. (2011).Healthcare-seeking behaviour for common infectious disease-related illnesses in rural Kenya: a community-based house-to-house survey.*Journal of health, population, and nutrition, 29*(1), 61.

Calle, M. (2006). Bayesian survival analysis modeling applied to sensory shelf life of foods. *Food quality and preference, 17*(3-4), 307-312.

Chatterjee, S. (2016). A study of epidemiological factors related to acute respiratory infection in under five children attending the immunization clinic of Calcutta National Medical College and Hospital. *Internet J Pulm Med, 7*(2), 1-6.

CSA. (2012). Central Statistical Agency, Addis Ababa, Ethiopia. Measure DHS, ICF Macro Calverton, Maryland, USA. Ethiopia Demographic and Health Survey.

CSA, I. (2016). Central statistical agency (CSA) [Ethiopia] and ICF. *Ethiopia demographic and health survey, Addis Ababa, Ethiopia and Calverton, Maryland, USA.*

Dadi, A.F., Kebede, Y. and Birhanu, Z. (2014). Determinants of Pneumonia in Children Aged Two Months to Five Years in Urban Areas of Oromia Zone, Amhara Region, Ethiopia.*Open Access Library Journal,* 1: e1044. http://dx.doi.org/10.4236/oalib.1101044.

Ekure, E. N., Esezobor, C. I., Balogun, M. R., Mukhtar-Yola, M., Ojo, O. O., Emodi, I. J., Ogala, W. N. (2013). Mothers and childhood pneumonia: what should the focus of public campaigns be? *Nigerian Journal of Paediatrics, 40*(1), 24-29.

Fahrmeir, L., Lang, S. (2001). Bayesian inference for generalized additive mixed models based on Markov random field priors.*Journal of the Royal Statistical Society: Series C (Applied Statistics), 50*(2), 201-220.

Fatmi, Z., White, F. (2002). A comparison of cough and cold and pneumonia: risk factors for pneumonia in children under 5 years revisited. *International Journal of Infectious Diseases, 6*(4), 294-301.

Federal Democratic Republic of Ethiopia Ministry of Health. Health Sector Transformation Plan, August 2015.PAGE 24.

Fekadu, G. A., Terefe, M. W., Alemie, G. A. (2014). Prevalence of pneumonia among under-five children in Este Town and the surrounding rural Kebeles, Northwest Ethiopia: a community based cross sectional study.*Science Journal of Public Health, 2*(3), 150-5.

Geberetsadik, A., Worku, A., Berhane, Y. (2015). Factors associated with acute respiratory infection in children under the age of 5 years: evidence from the 2011 Ethiopia Demographic and Health Survey. *Pediatric health, medicine and therapeutics, 6*, 9

Gelman A., Carlin J.B., Stern H.S., Rubin D.B. (2000). *Bayesian data analysis* ,Chapman Hall/CRC, USA.

Gelman, A., Hill, J. (2006).*Data analysis using regression and multilevel/hierarchical models.* Cambridge university press.

Gelman, A., Hwang, J. (2013). Understanding predictive information criteria for Bayesian models.

Gelman, A., B., C. J., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. (2014).*Bayesian data analysis.* New York, NY, USA: Chapman Hall.

Ghimire, M., Pradhan, Y. V., Maskey, M. K. (2010). Community-based interventions for diarrhoeal diseases and acute respiratory infections in Nepal.*Bulletin of the World Health Organization, 88,* 216-221.

Gilks, W. R., Richardson, S., Spiegelhalter, D. (1995).*Markov chain Monte Carlo in practice.* Chapman and Hall/CRC.

Gordvil-Merino, A., Gurdia-Olmos, J., Per-Cebollero, M., de la Fuente-Solanas, E. I. (2010). Classical and Bayesian estimation in the logistic regression model applied to diagnosis of child Attention Deficit Hyperactivity Disorder.*Psychological reports, 106*(2), 519-533.

Harerimana, J. M., Nyirazinyoye, L., Thomson, D. R., Ntaganira, J. (2016). Social, economic and environmental risk factors for acute lower respiratory infections among children under five years of age in Rwanda.*Archives of Public Health, 74*(1), 19.

Hug, L., Sharrow, D., You, D. (2017). Levels trends in child mortality: report 2017. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation.

Ibrahim J.G., Chen M.H. (2000). *Power Prior Distributions for Regression Models,* Statistical Science, 15 (1), 46-60.

International Vaccine Access Center (IVAC). Pneumonia and Diarrhea Progress Report 2016; Johns Hopkins Bloomberg School of Public Health: Baltimore, MD, USA, 2016.

Jabessa, S. (2015). Multilevel Analysis of Acute Respiratory Infection Symptoms among under Five Children in Ethiopia.*Journal of Biometrics Biostatistics, 6*(4), 1.

Jackson, S., Mathews, K. H., Pulani, D., Falconer, R., Rudan, I., Campbell, H., Nair, H. (2013). Risk factors for severe acute lower respiratory infections in childrena systematic review and meta-analysis.*Croatian medical journal, 54*(2), 110-121.

Kawo, K. N., Asfaw, Z. G., Yohannes, N. (2018). Multilevel Analysis of Determinants of Anemia Prevalence among Children Aged 659 Months in Ethiopia: Classical and Bayesian Approaches.*Anemia, 2018.*

Keter, P., Mbakaya, C., Gikunju, J., Mutai, J. (2015). Knowledge, Perceptions and Practices of Mothers in Relation to Childhood Pneumonia at Kapsabet District Hospital in Nandi County, Kenya. - Google Search. Retrieved August 16, 2016, from https://www.google.com/search?

Krishnan A, Amarchand R, Gupta V, Lafond KE, Suliankatchi RA, Saha S *et al.* (2015): Epidemiology of acute respiratory infections in children - preliminary results of a cohort in a rural north Indian community. BMC infectious diseases, 15: 462.

Liu L, Johnson HL, Cousens S, Perin J, Scott S, et al. (2012) Global, regional, and national causes of child mortality: An updated systematic analysis for 2010 with time trends since 2000. Lancet 379: 215161

Liu, M. M., Wang, D., Zhao, Y., Liu, Y. Q., Huang, M. M., Liu, Y., Dong, G. H. (2013). Effects of outdoor and indoor air pollution on respiratory health of Chinese children from 50 kindergartens. *Journal of epidemiology, 23*(4), 280-287.

Mahanta, J., Biswas, S. C., Roy, M. K., Islam, M. A. (2015). A comparison of Bayesian and classical approach for estimating Markov based logistic model. *American Journal of Mathematics and Statistics, 5*(4), 178-183.

Manu, G. (2014). Factors that influence acute respiratory infections among children below five years in Ghana.

Miller, N., Amouzou, A., Tafesse, M. amd Hazel E. (2014) Integrated Community Case management of common childhood illness in Ethiopia: implementation strength and quality of care. The American journal of tropical medicine and hygiene, 91(2), 424-434.

Mohammadi, M., Mohammadi, P. (2018). Indoor air pollution and acute respiratory infection among children: an update biomass smoke.*Journal of Air Pollution and*

Health, 3(1), 49-62.

Montasser, N., Helal, R., Rezq, R. (2012). Assessment and classification of acute respiratory tract infections among Egyptian rural children.*British Journal of Medicine and Medical Research, 2(2), 216.*

Mortal, G. B. D. (2015). Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.*Lancet, 385*(9963), 117-171.

Muthoni, K. A., Ngesa, O. O. (2017). Determinants and spatial modeling of acute respiratory infections (ARI) among children less than five years in Kenya. *American Journal of Theoretical and Applied Statistics, 6*(2), 123.

Nair, H., Simes, E. A., Rudan, I., Gessner, B. D., Azziz-Baumgartner, E., Zhang, J. S. F., Baggett, H. C. (2013). Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis.*The Lancet, 381*(9875), 1380-1390.

Noordam, A. C., Carvajal-Velez, L., Sharkey, A. B., Young, M., Cals, J. W. (2015). Care seeking behavior for children with suspected pneumonia in countries in sub-Saharan Africa with high pneumonia mortality.*PloS one, 10*(2), e0117919.

Ntzoufras, I. (2011).*Bayesian modeling using WinBUGS*(Vol. 698). John Wiley Sons.

R. J. Little and D. B. Rubin,*in Statistical Analysis with Missing Data,* vol. 88, pp. 125134, John Wiley, New York, NY, USA, 1987.

Ramani, V. K., Pattankar, J., Puttahonnappa, S. K. (2016). Acute respiratory infections among under-five age group children at urban slums of gulbarga city: A longitudinal study. *Journal of clinical and diagnostic research: JCDR, 10*(5), LC08.

Rasbash, J., Browne, W., Healy, M., Cameron, B., Charlton, C. (2005). MLwiN version 2.02 [Computer software].*Bristol, England: Centre for Multilevel Modelling.*

Rehman, A., Shaikh, B. T., Ronis, K. A. (2014). Health care seeking patterns and out of pocket payments for children under five years of age living in Katchi Abadis (slums), in Islamabad, Pakistan. *International Journal for Equity in Health, 13*(1), 1.

Richardson, S. a. (2003). Bayesian hierarchical models in ecological studies of health environment effects. Environ metrics. 129-147.

Sanbata, H., Asfaw, A., Kumie, A. (2014). Association of biomass fuel use with acute respiratory infections among under-five children in a slum urban of Addis Ababa, Ethiopia. *BMC public health, 14*(1), 1122.

Sainani, K. (2010). The Importance of Accounting for Correlated Observations. PM and R. https://doi.org/10.1016/j.pmrj.2010.07.482. 2(9), 858-861.

Salameh, P., Waked, M., Khayat, G., Dramaix, M. (2014). Bayesian and frequentist comparison for epidemiologists: A non-mathematical application on logistic regressions. *The Open Epidemiology Journal,* 7(1).

Schmidt, W. P., Cairn cross, S., Barreto, M. L., Close, T., Genser, B. (2009). Recent diarrhoeal illness and risk of lower respiratory infections in children under the age of 5 years. *International journal of epidemiology,* 38(3), 766-772.

Simoes, E. A., Cherian, T., Chow, J., Shahid-Salles, S. A., Laxminarayan, R., John, T. J. (2006). Acute respiratory infections in children. *In Disease Control Priorities in Developing Countries. 2nd edition.* The International Bank for Reconstruction and Development/the World Bank.

Snijders, T. and Bosker, R., (1999). Multilevel Analysis: an Introduction to Basic and Advanced Multilevel Modeling. London/ Thousand Oaks/ New Delhi: Sage Publications.

Sonego, M., Pellegrin, M. C., Becker, G., Lazzerini, M. (2015). Risk factors for mortality from acute lower respiratory infections (ALRI) in children under five years of age in low and middle-income countries: a systematic review and meta-analysis of observational studies. *PloS one, 10(1), e0116380.*

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology), 64*(4), 583-639.

Spiegelhalter, D. J., Abrams, K. R., Myles, J. P. (2004). *Bayesian approaches to clinical trials and health-care evaluation* (Vol. 13). John Wiley Sons.

Survey of the last ten kilometers project (L10K), (2013). Integrated Community Case Management (iCCM) Survey in Amhara, SNNP and Tigray regions.

Taylor, E. T., Nakai, S. (2012). Prevalence of acute respiratory infections in women and children in Western Sierra Leone due to smoke from wood and charcoal stoves. *International journal of environmental research and public health, 9*(6), 2252-2265.

Ujunwa, F. A., Ezeonu, C. T. (2014). Risk Factors for Acute Respiratory Tract Infections in Under five Children in Enugu Southeast Nigeria. *Annals of medical and health sciences research, 4*(1), 95-99.

UNICEF (2014). Pneumonia and Diarrhea Progress Report. International Vaccine Access Center (IVAC).Johns Hopkins Bloomberg, School of Public Health, www.jhsph.edu/ivac.

UNICEF, (2015). Committing to child survival: a promise renewed progress report 2014.

United Nations Childrens Fund (UNICEF). Committing to Child Survival: A Promise Renewed Progress Report 2015.

UNICEF (2016). A Fair Chance for Every Child: The State of the Worlds Children 2016.

Walke, S. P., Das, R., Acharya, A. S., Pemde, H. K. (2014). Incidence, pattern, and severity of acute respiratory infections among infants and toddlers of a peri-urban area of Delhi: a 12-month prospective study. *International scholarly research notices.*

Walker, C. L. F., Perin, J., Katz, J., Tielsch, J. M., Black, R. E. (2013). Diarrhea as a risk factor for acute lower respiratory tract infections among young children in low income settings. *Journal of global health, 3*(1).

Williams BG, Gouws E, Boschi PC, Bryce J, Dye C (2002) Estimates of Worldwide distribution of child death from acute respiratory infections. Lancet infec Dis 2(1): 25-32.

World Health Organization (2011). Initiative for vaccine research.

World Health Organization. (2017). World Health Statistics 2017: monitoring health for the Sustainable Development Goals. Geneva: World Health Organization; 2017.

WHO. (2016). Pneumonia Fact Sheet. Accessed on 17 October 2016. Retrieved from http://www. who.int/mediacentre/factsheets/fs331/en/index.html

World Health Organization. (2018). Global Nutrition Targets 2025 Breastfeeding Policy Brief.

WLF, World lung foundation Report 2011, Geneva, World health organization.

Appendices

Appendix A: Result of Chi-square Test.

Table 5.1: Chi-square association between ARI status and predictor variables

Variable names	Chi-square value	p-value
Sex of child	0.0412	0.839
Child age	78.1217	0.000*
Mothers age	0.6024	0.740
Wealth index	29.140	0.000*
Place of residence	14.3655	0.000*
Maternal educations	24.7906	0.000*
Maternal occupations	3.9042	0.048*
Vitamin A supplement	4.508	0.034*
Breastfeeding	54.3091	0.000*
History of diarrhea	421.8497	0.000*
Type of fuel	6.8230	0.009*
Wasting	4.1869	0.041*
Stunting	33.4323	0.000*
Number of children	2.3503	0.309
Source of water	9.253	0.002*
Region	249.0584	0.000*

* indicates a significant variable at 5% level of significance

Appendix B: Result of Bayesian multilevel Logistic regression Analysis of MLwiN output.

Table 5.2: Random Intercept Bayesian multilevel model.

Fixed effect							
Variables	Category	Coefficient	SD	MC error	2.5%	50%	97.5%
-	Intercept	-3.280	0.367	0.0063	-4.035	-3.270	-2.563
Sex of child	Male(ref)						
	Female	0.007	0.080	0.0003	-0.151	0.008	0.170
Child age	<6 (ref)						
	6-11	0.245	0.162	0.0007	-0.066	0.243	0.563
	12-23	0.087	0.147	0.0007	-0.195	0.088	0.369
	24-35	-0.177	0.169	0.0009	-0.509	-0.180	0.153
	36-47	-0.193	0.186	0.001	-0.556	-0.193	0.176
	48-59	-0.539	0.196	0.001	-0.925	-0.539	-0.148
Maternal age	15-19(ref)						
	20-34	0.153	0.223	0.0022	-0.264	0.146	0.626
	35-49	0.050	0.249	0.0021	-0.419	0.043	0.556
Wealth index of household	Poorest (ref)						
	Poorer	-0.214	0.124	0.0005	-0.459	-0.214	0.024
	Middle	-0.609	0.145	0.0005	-0.892	-0.609	-0.331
	Richer	-0.812	0.161	0.0006	-1.130	-0.813	-0.509
	Richest	-0.598	0.204	0.0009	-0.994	-0.595	-0.213
Place of residence	Rural (ref)						
	Urban	0.244	0.220	0.001	-0.194	0.245	0.684
Breast feeding	Never bf(ref)						
	Ever bf, not currently	0.229	0.253	0.0025	-0.249	0.219	0.737
	Still bf	0.356	0.258	0.0028	-0.150	0.349	0.868

Maternal education	No education(ref)						
	Primary	0.154	0.105	0.0004	-0.051	0.154	0.362
	Secondary and higher	-0.449	0.201	0.0007	-0.851	-0.443	-0.065
Vitamin A Supplement	No (ref)						
	Yes	-0.183	0.088	0.0003	-0.355	-0.184	-0.010
Had diarrhea	No(ref)						
	Yes	1.435	0.096	0.0003	1.245	1.590	1.621
Fuel type	Unclean/unsafe(ref)						
	Safe/clean	-0.065	0.266	0.0009	-0.587	-0.061	0.454
Maternal Work	Not working(ref)						
	Working	0.268	0.094	0.0003	0.084	0.268	0.459
Wasting	Not wasted (ref)						
	Wasted	0.122	0.120	0.0004	-0.116	0.119	0.354
Stunting	Not stunted(ref)						
	Stunted	0.456	0.087	0.0003	0.292	0.456	0.632
Number of living child	1-3 (ref)						
	4-6	0.126	0.100	0.0004	-0.070	0.123	0.319
	Above 6	0.212	0.150	0.0006	-0.088	0.213	0.499
Source of drinking water	Unprotected(ref)						
	Protected	-0.203	0.095	0.0003	-0.387	-0.202	-0.016
Random effect							
Var(U$_{0j}$) = σ^2_{u0}		1.107	0.158	0.0007	0.829	1.095	1.456

Note:

Ref=Referance category

σ^2_{u0}=regions variance

Table 5.3: Random coefficient Bayesian multilevel model.

Fixed effects								
Variables	category	Coefficient	MC error	SD	Exp(β)	2.5%	50%	97.5%
–	Intercept	-3.305	0.0057	0.353	0.0367	-4.025	-3.307	-2.615
Sex of child	Male(ref)							
	Female	0.011	0.0003	0.081	1.011	-0.143	0.012	0.169
Child age	<6 (ref)							
	6-11	0.259	0.0008	0.163	1.295	-0.061	0.263	0.579
	12-23	0.095	0.0008	0.151	1.099	-0.199	0.094	0.388
	24-35	-0.155	0.001	0.175	0.856	-0.489	-0.157	0.188
	36-47	-0.164	0.0011	0.190	0.849	-0.541	-0.159	0.192
	48-59	-0.514	0.0011	0.201	0.598	-0.908	-0.517	-0.130
Maternal age	15-19(ref)							
	20-34	0.142	0.0023	0.221	1.153	-0.282	0.139	0.583
	35-49	0.029	0.002	0.251	1.029	-0.462	0.024	0.532
Wealth index	Poorest (ref)							
	Poorer	-0.222	0.0005	0.125	0.800	-0.472	-0.224	0.021
	Middle	-0.632	0.0006	0.147	0.532	-0.927	-0.629	-0.343
	Richer	-0.877	0.0015	0.245	0.416	-1.417	-0.863	-0.435
	Richest	-0.638	0.001	0.210	0.528	-1.051	-0.635	-0.226
Place of residence	Rural (ref)							
	Urban	0.264	0.0011	0.225	1.302	-0.174	0.262	0.709
Breast feeding	Never bf (ref)							
	Ever bf, not currently	0.188	0.0024	0.250	1.207	-0.273	0.183	0.697
	Still breast feed	0.337	0.0028	0.258	1.400	-0.158	0.333	0.870

Maternal educ.	No education(ref)							
	Primary	0.140	0.0004	0.106	1.150	-0.064	0.140	0.350
	Secondary and higher	-0.454	0.0007	0.205	0.635	-0.875	-0.459	-0.067
Vitamin A	No (ref)							
	Yes	-0.182	0.0003	0.088	0.834	-0.359	-0.183	-0.011
Had diarrhea	No(ref)							
	Yes	1.455	0.0003	0.099	4.284	1.260	1.455	1.649
Fuel type	Unclean/unsafe(ref)							
	Safe/clean	-0.062	0.001	0.269	0.939	-0.607	-0.051	0.445
Maternal work	Not working(ref)							
	Working	0.270	0.0007	0.137	1.309	0.011	0.271	0.540
Wasting	Not wasted(ref)							
	Wasted	0.125	0.0004	0.122	1.133	-0.117	0.125	0.360
Stunting	Not stunted(ref)							
	Stunted	0.467	0.0003	0.089	1.595	0.291	0.464	0.640
Number of living child	1-3 (ref)							
	4-6	0.126	0.0004	0.101	1.134	-0.068	0.121	0.328
	above 6	0.217	0.0006	0.152	1.242	-0.072	0.215	0.515
Source of drinking water	Unprotected(ref)							
	Protected	-0.206	0.0004	0.098	0.814	-0.400	-0.204	-0.016
Random effect								
σ^2_{u0}		1.295	0.0012	0.203		0.925	1.284	1.724
Random slope								
σ^2_{u11}		0.725	0.0049	0.351		0.223	0.666	1.606
σ^2_{u21}		0.502	0.0032	0.223		0.188	0.468	1.031

Note:

Ref=Referance category

σ^2_{u0}=regions variance

σ^2_{u11}=wealth index variance

σ^2_{u21}=maternal working variance

Appendix C: MLwiN Result for Equations of Bayesian multi-level logistic regression models.

MLwiN equation result for Bayesian multilevel logistic regression-empty model.

$$\text{ARI status}_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{cons}$$

$$\beta_{0j} = -2.789(0.072) + u_{0j}$$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[0.998(0.136)\right]$$

$$\text{var}(\text{ARI status}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}$$

PRIOR SPECIFICATIONS

$$p(\beta_0)\ \alpha\ 1$$

$$p(1/\sigma^2_{u0}) \sim \text{Gamma}(0.001,0.001)$$

Deviance(MCMC) = 5097.643(9918 of 9918 cases in use)

MLwiN equation result for Bayesian multilevel logistic regression-random intercept model

$\text{ARI status}_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$

$$\begin{aligned}\text{logit}(\pi_{ij}) = {} & \beta_{0j}\text{cons} + 0.007(0.080)\text{Female}_{ij} + 0.245(0.162)\text{6-11 months}_{ij} + 0.087(0.147)\text{12-23 months}_{ij} + -0.177(0.169)\text{24-35 months}_{ij} + \\ & -0.193(0.186)\text{36-47 months}_{ij} + -0.539(0.196)\text{48-59 months}_{ij} + 0.153(0.223)\text{20-34}_{ij} + 0.050(0.249)\text{35-49}_{ij} + -0.214(0.124)\text{Poorer}_{ij} + \\ & -0.609(0.145)\text{Middle}_{ij} + -0.812(0.161)\text{Richer}_{ij} + -0.598(0.204)\text{Richest}_{ij} + 0.244(0.220)\text{Urban}_{ij} + \\ & 0.229(0.253)\text{Ever breast feed,not currently}_{ij} + 0.356(0.258)\text{Still breast feed}_{ij} + 0.154(0.105)\text{Primary}_{ij} + \\ & -0.449(0.201)\text{Secondary and above}_{ij} + -0.183(0.088)\text{Have vitamin A}_{ij} + 1.435(0.096)\text{Had diarrhea}_{ij} + -0.065(0.266)\text{safe/clean}_{ij} + \\ & 0.268(0.094)\text{Working}_{ij} + 0.122(0.120)\text{Wasted}_{ij} + 0.456(0.087)\text{Stunted}_{ij} + 0.126(0.100)\text{4-6 child}_{ij} + 0.212(0.150)\text{6 and above}_{ij} + \\ & -0.203(0.095)\text{Protected/improved}_{ij}\end{aligned}$$

$\beta_{0j} = -3.280(0.367) + u_{0j}$

$$\left[u_{0j}\right] \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \left[1.107(0.158)\right]$$

$\text{var}(\text{ARI status}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}$

PRIOR SPECIFICATIONS

$p(\beta_0)\ \alpha\ 1$

$p(\beta_1)\ \alpha\ 1$

$p(\beta_2)\ \alpha\ 1$

$p(\beta_3)\ \alpha\ 1$

$p(\beta_4)\ \alpha\ 1$

$p(\beta_5)\ \alpha\ 1$

$p(\beta_6)\ \alpha\ 1$

$p(\beta_7)\ \alpha\ 1$

$p(\beta_8)\ \alpha\ 1$

$p(\beta_9)\ \alpha\ 1$

$p(\beta_{10})\ \alpha\ 1$

$p(\beta_{11})\ \alpha\ 1$

$p(\beta_{12})\ \alpha\ 1$

$p(\beta_{13})\ \alpha\ 1$

$p(\beta_{14})\ \alpha\ 1$

$p(\beta_{15})\ \alpha\ 1$

$p(\beta_{16})\ \alpha\ 1$

$p(\beta_{17})\ \alpha\ 1$

$p(\beta_{18})\ \alpha\ 1$

$p(\beta_{19})\ \alpha\ 1$

$p(\beta_{20})\ \alpha\ 1$

$p(\beta_{21})\ \alpha\ 1$

$p(\beta_{22})\ \alpha\ 1$

$p(\beta_{23})\ \alpha\ 1$

$p(\beta_{24})\ \alpha\ 1$

$p(\beta_{25})\ \alpha\ 1$

$p(\beta_{26})\ \alpha\ 1$

$p(1/\sigma^2_{u0}) \sim \text{Gamma}(0.001,0.001)$

Deviance(MCMC) = 4699.415(9918 of 9918 cases in use)

MLwiN equation result for Bayesian multilevel logistic regression-random coefficient model.

$\text{ARI status}_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$

$\text{logit}(\pi_{ij}) = \beta_{0j}\text{cons} + 0.011(0.081)\text{Female}_{ij} + 0.259(0.163)\text{6-11 months}_{ij} + 0.095(0.151)\text{12-23 months}_{ij} + -0.155(0.175)\text{24-35 months}_{ij} + -0.164(0.190)\text{36-47 months}_{ij} + -0.514(0.201)\text{48-59 months}_{ij} + 0.142(0.221)\text{20-34}_{ij} + 0.029(0.251)\text{35-49}_{ij} + -0.222(0.125)\text{Poorer}_{ij} + -0.632(0.147)\text{Middle}_{ij} + \beta_{1j}\text{Richer}_{ij} + -0.638(0.210)\text{Richest}_{ij} + 0.264(0.225)\text{Urban}_{ij} + 0.188(0.250)\text{Ever breast feed,not currently}_{ij} + 0.337(0.258)\text{Still breast feed}_{ij} + 0.140(0.106)\text{Primary}_{ij} + -0.454(0.205)\text{Secondary and above}_{ij} + -0.182(0.088)\text{Have vitamin A}_{ij} + 1.455(0.099)\text{Had diarrhea}_{ij} + -0.062(0.269)\text{safe/clean}_{ij} + \beta_{2j}\text{Working}_{ij} + 0.125(0.122)\text{Wasted}_{ij} + 0.467(0.089)\text{Stunted}_{ij} + 0.126(0.101)\text{4-6 child}_{ij} + 0.217(0.152)\text{6 and above}_{ij} + -0.206(0.098)\text{Protected/improved}_{ij}$

$\beta_{0j} = -3.305(0.353) + u_{0j}$

$\beta_{1j} = -0.877(0.245) + u_{1j}$

$\beta_{2j} = 0.270(0.137) + u_{2j}$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u)\ :\ \Omega_u = \begin{bmatrix} 1.295(0.203) & & \\ -0.177(0.238) & 0.725(0.351) & \\ -0.237(0.168) & -0.148(0.191) & 0.502(0.223) \end{bmatrix}$$

$\text{var}(\text{ARI status}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}$

PRIOR SPECIFICATIONS

$p(\beta_0)\ \alpha\ 1$
$p(\beta_1)\ \alpha\ 1$
$p(\beta_2)\ \alpha\ 1$
$p(\beta_3)\ \alpha\ 1$
$p(\beta_4)\ \alpha\ 1$
$p(\beta_5)\ \alpha\ 1$
$p(\beta_6)\ \alpha\ 1$
$p(\beta_7)\ \alpha\ 1$
$p(\beta_8)\ \alpha\ 1$
$p(\beta_9)\ \alpha\ 1$
$p(\beta_{10})\ \alpha\ 1$
$p(\beta_{11})\ \alpha\ 1$
$p(\beta_{12})\ \alpha\ 1$
$p(\beta_{13})\ \alpha\ 1$
$p(\beta_{14})\ \alpha\ 1$
$p(\beta_{15})\ \alpha\ 1$
$p(\beta_{16})\ \alpha\ 1$
$p(\beta_{17})\ \alpha\ 1$
$p(\beta_{18})\ \alpha\ 1$
$p(\beta_{19})\ \alpha\ 1$
$p(\beta_{20})\ \alpha\ 1$
$p(\beta_{21})\ \alpha\ 1$
$p(\beta_{22})\ \alpha\ 1$
$p(\beta_{23})\ \alpha\ 1$
$p(\beta_{24})\ \alpha\ 1$
$p(\beta_{25})\ \alpha\ 1$
$p(\beta_{26})\ \alpha\ 1$

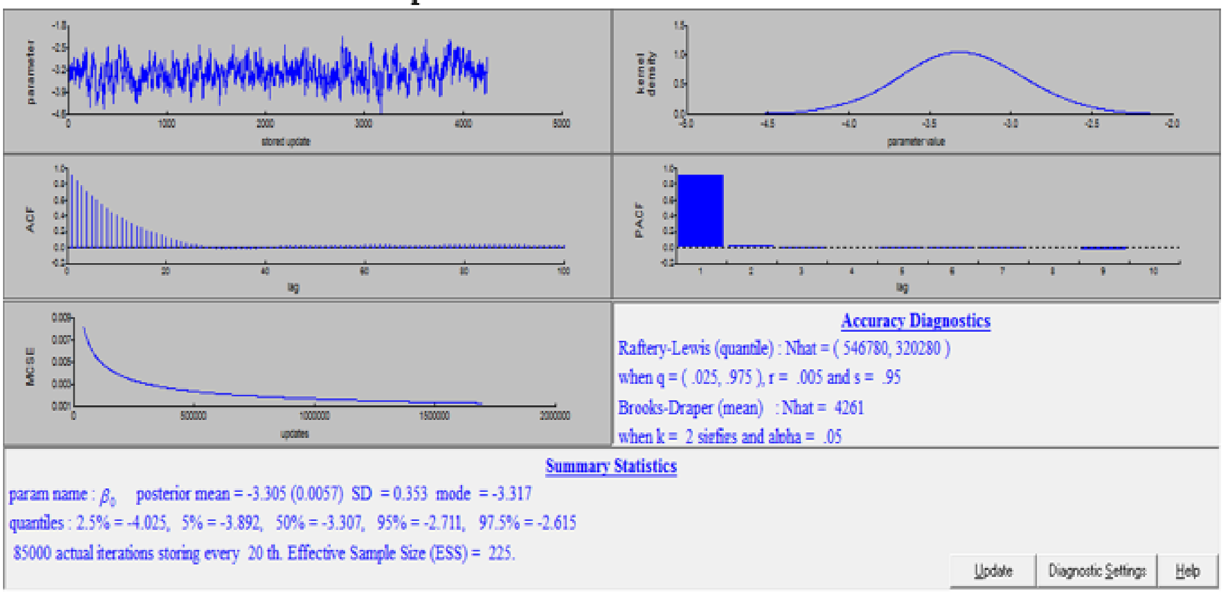
$$p(\Omega_u) \sim \text{inverse Wishart}_3[\,3{*}S_u, 3],\ \ S_u = \begin{bmatrix} 1.295 & & \\ -0.177 & 0.725 & \\ -0.237 & -0.148 & 0.502 \end{bmatrix}$$

Deviance(MCMC) = 4612.548(9918 of 9918 cases in use)

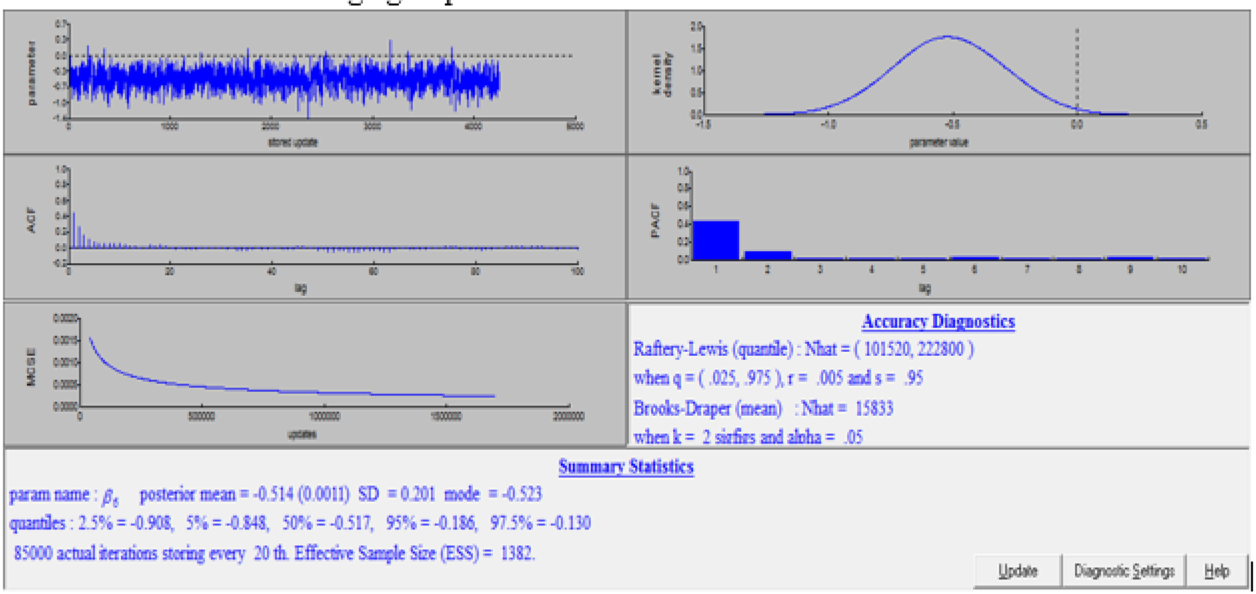
Appendix D: List of Figures for diagnostics

Figure 5.1: Plots of Bayesian Multilevel random Coefficients' Convergence Test

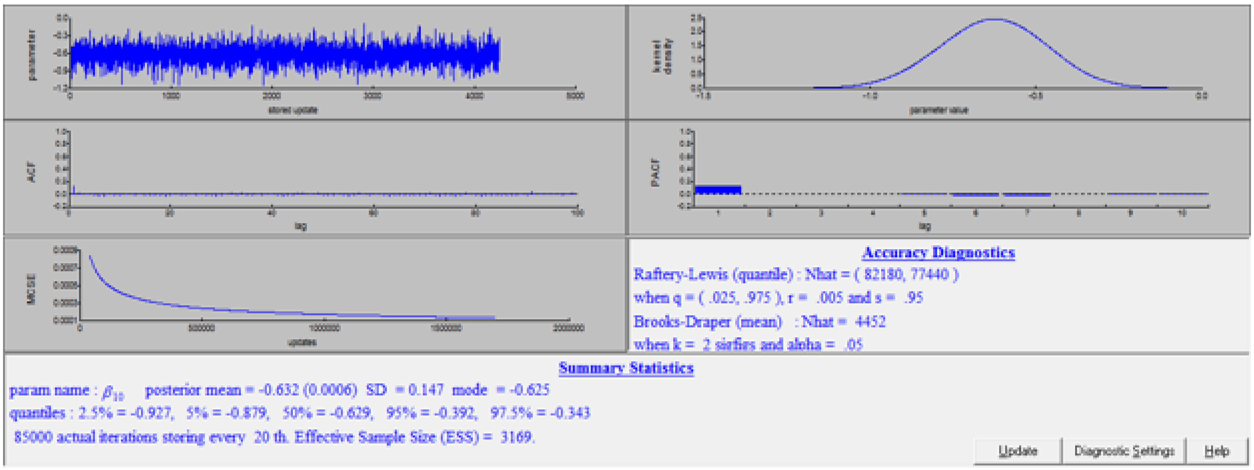
Intercept



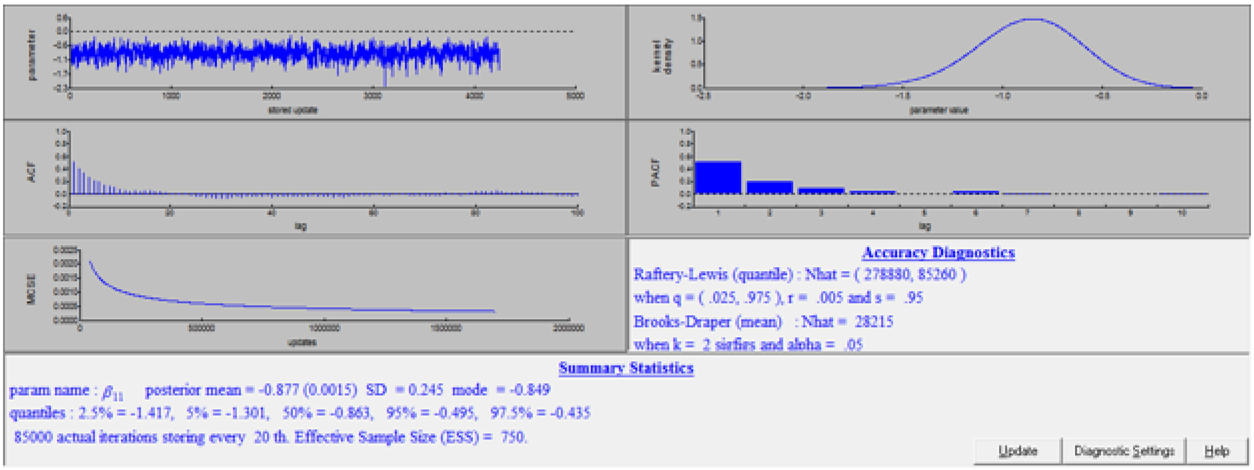
Coefficient of child age group in 48-59 months



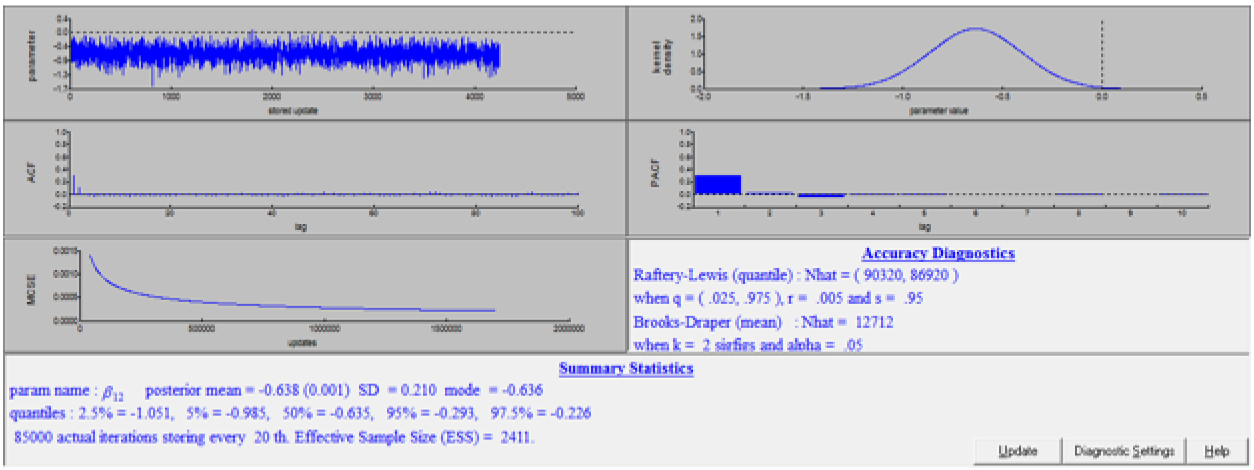
Coefficient of wealth index household in middle class



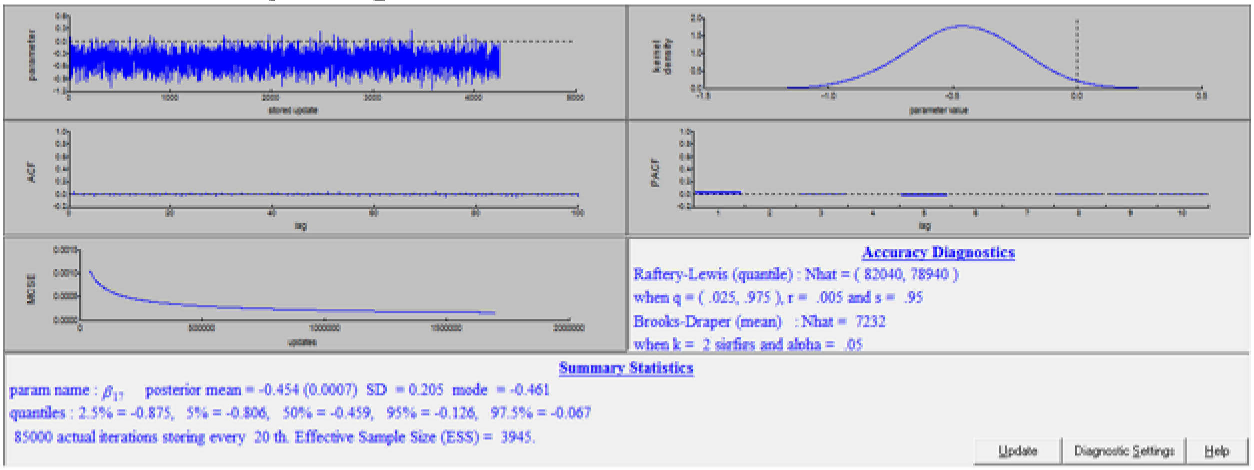
Coefficient of wealth index household in richer class



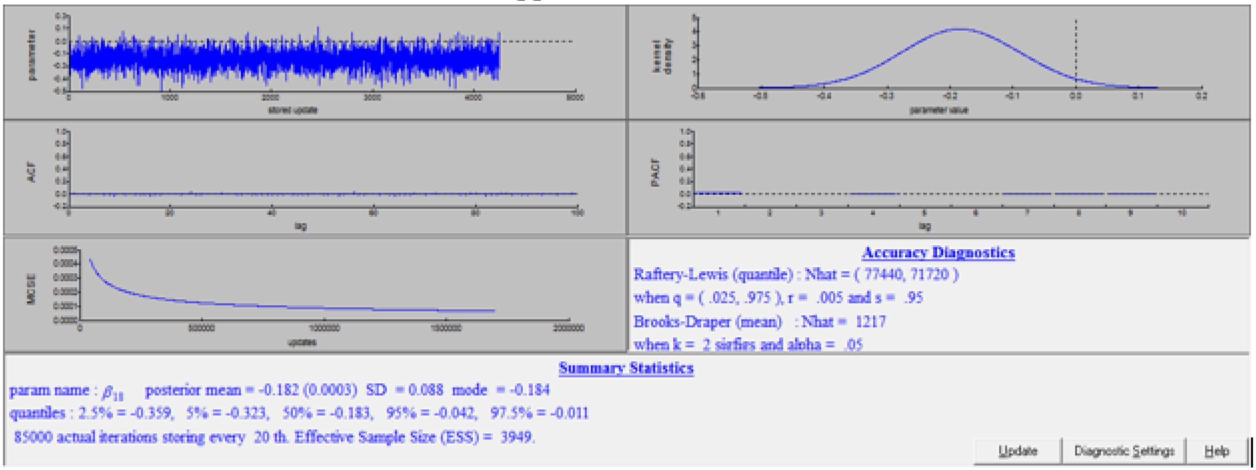
Coefficient of wealth index household in richest class



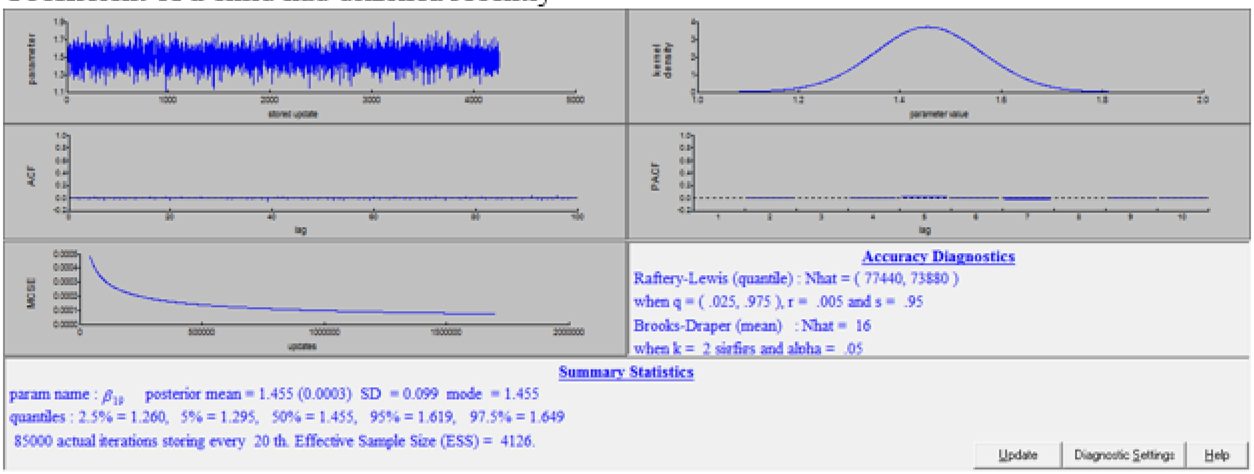
Coefficient secondary and higher education level



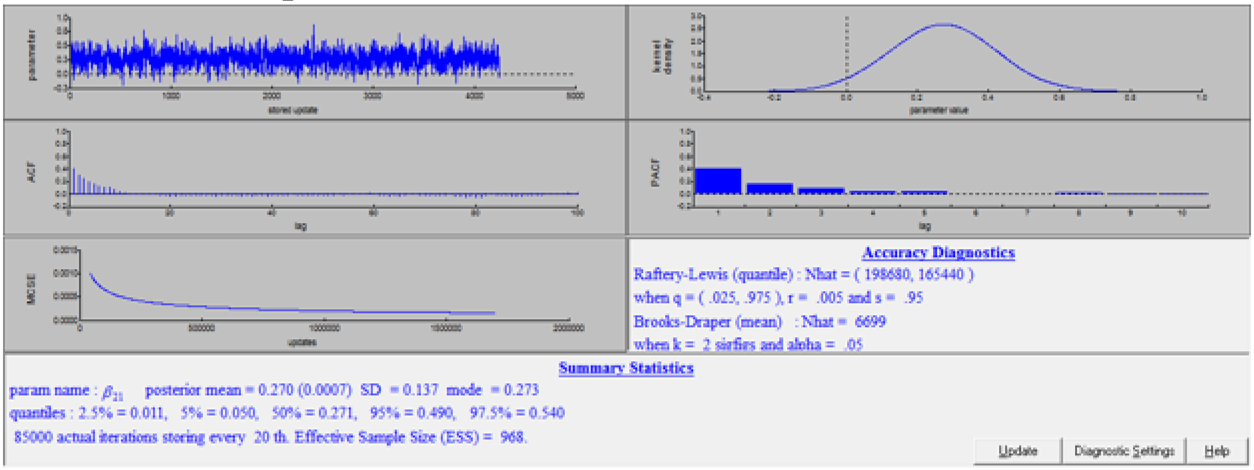
Coefficient of a child have vitamin A supplement



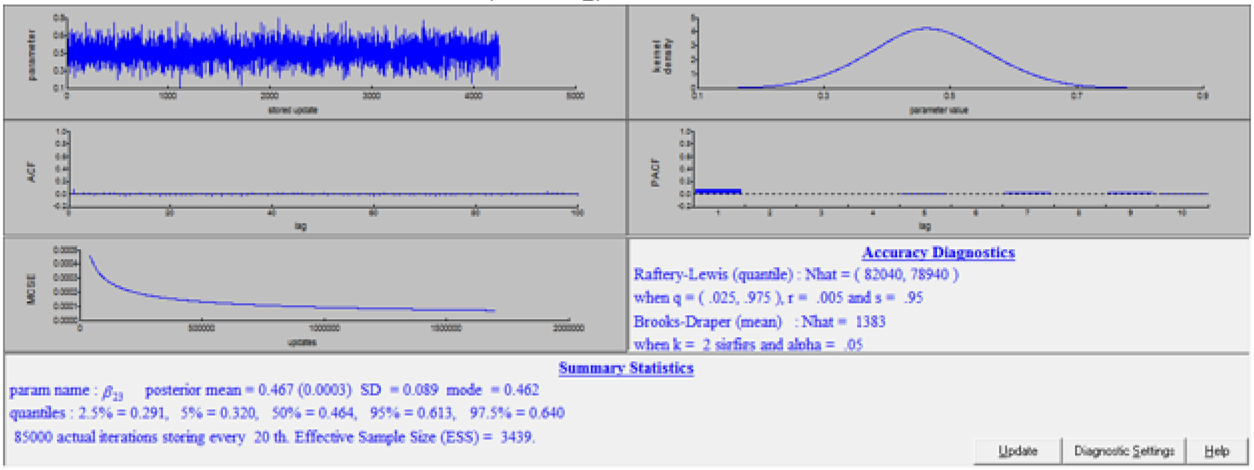
Coefficient of a child had diarrhea recently



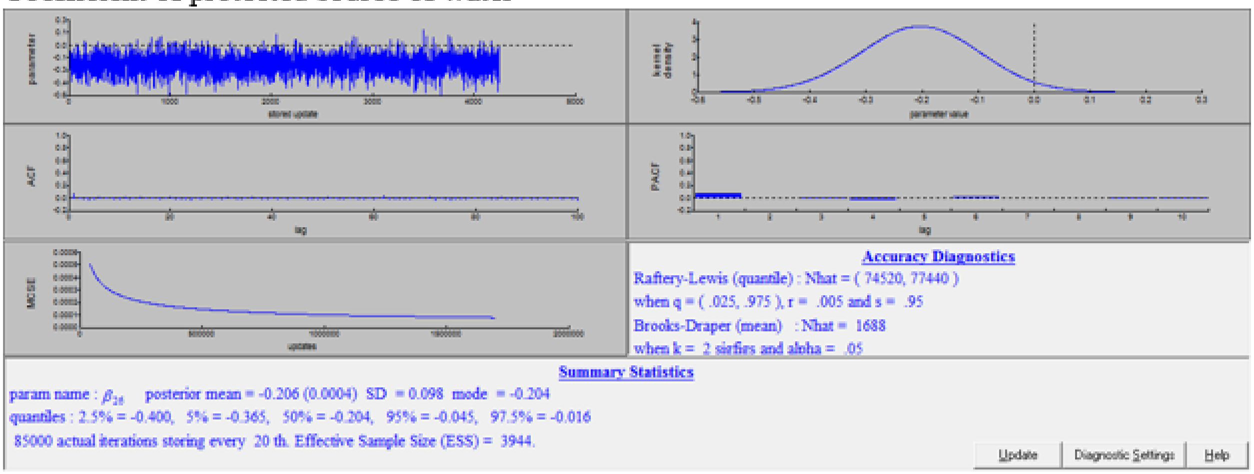
Coefficient of working mothers



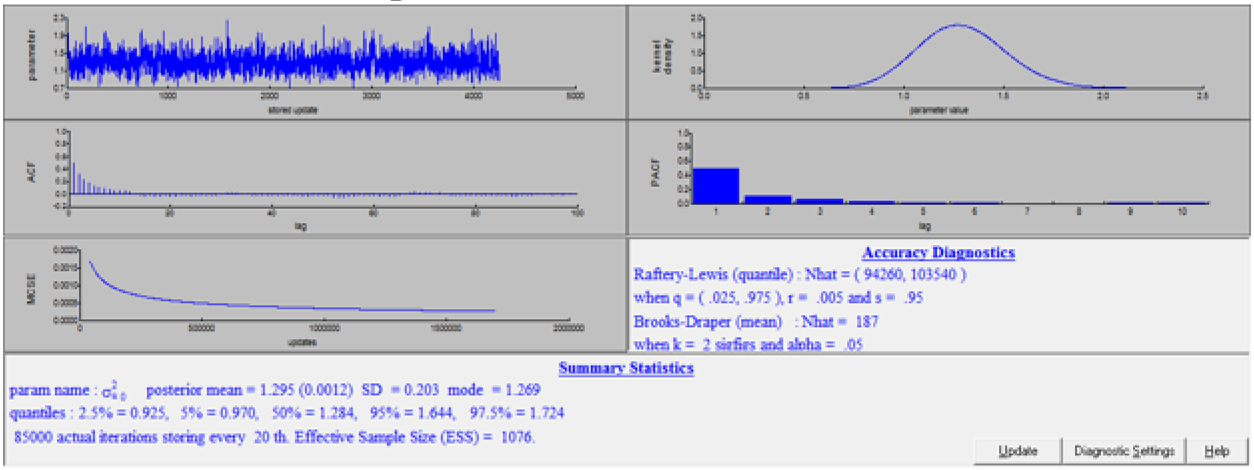
Coefficient of chronic malnutrition's (stunting)



Coefficient of protected source of water



Random Effect Variance of region



Random Effect Variance of working mothers