Determinants and Spatial Modeling of Acute Respiratory Infections (ARI) Among Children Less Than Five Years in Kenya

Kinyua Ann Muthoni*, Oscar Owino Ngesa

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email address: muthonikinyua12@gmail.com (K. A. Muthoni)
*Corresponding author


Received: February 20, 2017; Accepted: March 1, 2017; Published: March 21, 2017

Abstract: Bayesian disease mapping is a field of statistics that is used to model the spatial distribution of disease outcomes especially in application to studies in spatial biostatistics and also as a tool to help develop the required intervention strategies. In this study, we perform a spatial modeling of ARI among children less than five years in Kenya using data from the 2014 Kenya Demographic and Health Survey (KDHS). Four models were used in this study namely the logistic regression model, the normal unstructured heterogeneity random effects model, ICAR (Integrated Conditional Autoregressive) spatial random effects model and the convolution model. A full Bayesian approach was used and the models were implemented using the Winbugs software version 1.4. Model selection was based on the DIC value where the model with the lowest DIC value was considered to be the best. The convolution model was the best model in this case and was used to map ARI across the different counties in Kenya. The national prevalence was 47.3%. The prevalence was found to be highest in the counties in the western part of Kenya. From the analysis, it’s clear that ARI is still a menace that need to be controlled. Proper planning and allocation of resources need to be put in place by the county governments in order to curb the rising cases of ARI.

Keywords: Prevalence, Spatial, Bayesian

1. Introduction

ARI is an infection that interferes with normal breathing. It is classified based on the site of infection [1] as Acute Upper Respiratory Infections (AURI) and Acute Lower Respiratory Infections (ALRI). Acute Respiratory Infections (ARIs) are the leading cause of mortality in children under five years of age worldwide. Out of nearly 15 million children under 5 dying each year, 4 million die of ARIs. More than 90% of all these deaths occur in developing countries where children under 5 represent about 15% of the total population [2]. The World Health Organization estimates that the annual number of ARI-related deaths in children less than five years was 2.1 million in 2001 accounting for about 20% of all childhood deaths [3]. In Kenya, 19% of all the cases seen in outpatient clinics and hospitals are acute respiratory infection and mostly in the urban communities [4]. Children with acute respiratory infection, will at least have the following signs; cough, runny nose, fast breathing, and difficulty in breathing and chest in drawing. According to the 2014 Kenya Demographic and Health Survey (KDHS), the under-five mortality is 52 deaths per 1,000 live births. This means that 1 in 20 children dies before their fifth birthday. This is less than half the under-five mortality rate published in the 2003 KDHS when more than 2 in 20 children died of ARIs. Acute Upper Respiratory Infections include nasopharyngitis (commonly known as a cold), pharyngotonsillitis (inflammation of the tonsils) and otitis (inflammation of the ear). ALRI includes epiglottitis (is an inflammation of the epiglottis, laryngitis (inflammation of the larynx), laryngotracheitis (inflammation of the voice box), bronchitis, bronchiolitis and pneumonia. Upper Respiratory Infections are the most common infectious disease. The most common Lower Respiratory Infections (LRIs) in children are
pneumonia and bronchiolitis. Of the 6.3 million children who died before the age of years in 2013, pneumonia accounted for 0.953 million [14.9%] [5]. In Kenya, it is the second leading cause of mortality, accounting for more than 30,000 deaths in this age group annually [6]. It can be caused by viruses, bacteria or fungi. Pneumonia can be prevented by immunization, adequate nutrition and by addressing environmental factors. Pneumonia caused by bacteria can be treated with antibiotics, but only one third of children with pneumonia receive the antibiotics they need.

When we say that 19% of all the cases seen in outpatient clinics and hospitals in whole country are acute respiratory infection it does not necessarily mean that this number is shared equally by all the forty seven counties in Kenya. Several studies on acute respiratory infections for children under 5 in Kenya have been done previously [4] and [11]. However almost none of them has tried to identify the spatial effect on mortality on the specific counties. There may be a possibility that some counties are doing better than others. This means that the estimates discussed for the country could lead to disparities between counties. Different counties have different needs and challenges. Studying the geographical variation of neonatal mortality is of particular interest because access to antenatal or reproductive care vary and there exist regional differences in availability of services. A challenge in one county may not necessarily be a challenge in another county. Thus, generalizing ARIs at national level could lead to disparities between counties.

Moreover, Kenya is a multicultural country with diverse religious and cultural beliefs. The climatic conditions vary from one region to another. We always hear of marginalized areas. The budgetary priorities of each county depends on the needs of the people of that county. Understanding of geographical patterns of ARI can help county governments to be more accurate and specific when making resource allocations. Thus, if we went deeper to investigate each county separately the information would be much more helpful. Knowledge of what exists in different parts of any country is very important to the government for planning, evaluation, monitoring and execution of development projects aimed at improving the socio-economic conditions of our counties. Spatial modeling can be used to handle this. In this case the spatial units are the forty seven counties of Kenya. Analysis of data per county (lattice data) will be done in order to cater for these discrepancies.

Disparities in ARI cases across broad socioeconomic status and geographical regions have been reported internationally [7]. As earlier noted by [8] correct home based management is deficient and knowledge of danger symptoms was low. Within Kenya, there are disparities in ARI outcomes with respect to geographic regions and socioeconomic status. Understanding disparities in broad areas like for a whole country, while useful, is not likely to accurately reflect the heterogeneity in outcomes at the county level. Efforts to monitor and reduce ARI cases disparities can benefit greatly from quantifying variation across populations in small geographical areas like counties. An understanding of the geographic patterns of ARI can assist in improving health decision-making by health care planners like county governments to be more accurate and effective, for example by targeting policy development and resource allocation at areas of greater need [7]. This study was also carried out with the aim of establishing the distribution of prevalence of ARI across all the counties in Kenya. The use of disease maps to help in decision making in epidemiological and medical research is well recognized [7] and has been used in this study to map ARI in Kenya.

2. Methods

2.1. The Data

The data used in this study was obtained from the 2014 Kenya Demographic and Health Survey (KDHS). The sample for the 2014 KDHS was drawn from a master sampling frame, the Fifth National Sample Survey and Evaluation Program (NASSEP V). This is a frame that the Kenya National Bureau of Statistics (KNBS) currently operates to conduct household-based surveys throughout Kenya. Development of the frame began in 2012, and it contains a total of 5,360 clusters split into four equal subsamples. These clusters were drawn with a stratified probability proportional to size sampling methodology from 96,251 enumeration areas (EAs) in the 2009 Kenya Population and Housing Census. The 2014 KDHS used two subsamples of the NASSEP V frame that were developed in 2013. Approximately half of the clusters in these two subsamples were updated between November 2013 and September 2014. Kenya is divided into 47 counties that serve as devolved units of administration, created in the new constitution of 2010.

During the development of the NASSEP V, each of the 47 counties was stratified into urban and rural strata; since Nairobi County and Mombasa County have only urban areas, the resulting total was 92 sampling strata. A total of 39,679 households were selected for the sample. Of these households, 36,430 were successfully interviewed, yielding an overall household response rate of 99%.

2.2. Ethical Considerations

This study was based on secondary data with all participant identifiers removed. Survey procedures and instruments were approved by the Scientific and Ethical Review Committee of the Kenya Medical Research Institute (KEMRI) and by the Ethics Committee of the Opinion Research Corporation, Macro International Incorporated (ORC Macro Inc.), Calverton, USA. Ethical permission for use of the data in the present study was obtained from ORC Macro Inc. Details concerning the data collection protocols are available on the Measures Demographic and Health Surveys (DHS) website (http://dhsprogram.com/).

2.3. The Response Variable

The response variable was presence or absence of ARI in a child who is under five years of age which was coded with a
value zero to indicate absence of ARI and one to indicate presence of ARI. 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 preceding the survey.

2.4. Bivariate Data Analysis

Bivariate data analysis was carried out using Stata. The odds ratio estimates of the logistic regression model were employed to determine the significant risk factors of acute respiratory infection among children less than five years. Place of residence, wealth quantile and household size were found to be significantly associated with ARI because the 95% confidence interval for the odds ratio of these variables did not contain 1 while mother’s age, mother’s education level and type of cooking fuel were found to be insignificant (their 95% confidence interval for the odds ratio contained 1) and were therefore dropped from the analysis.

2.5. Model Specification

In this section, we shall look at some of the frequently encountered models for modeling the Bernoulli data at hand.

\[
\text{logit}(p_{ij}) = X^T \beta + v_i, \quad \text{Normal unstructured heterogeneity (UH) random effects model} \tag{3}
\]

In the third model, we added the structured random effects to the logistic regression model to get,

\[
\text{logit}(p_{ij}) = X^T \beta + u_i, \quad \text{ICAR spatial random effects model.} \tag{4}
\]

In model four the covariates and random effects are introduced as follows

\[
\text{logit}(p_{ij}) = X^T \beta + u_i + v_i, \quad \text{Convolution model.} \tag{5}
\]

The random effects \(u\) and \(v\) will be modeled using conditional autoregressive priors and normal distribution respectively. In this study, fully Bayesian inference will be used basing on the posterior distributions of the model parameters, which will be implemented by drawing random samples via Markov Chain Monte Carlo (MCMC) simulation techniques to all the models fitted.

2.6. Parameter Estimation

Model estimation was carried out using the Bayesian approach and appropriate prior distributions specified for all parameters of the models. In addition to the priors given to the random effects shown in the models above, non-informative priors were assigned to the regression coefficients. This was achieved by setting the covariate coefficients to have a highly dispersed normal distribution priors e.g

\[
p(\beta) \sim N(0,10000) \tag{6}
\]

Bayesian inference was used in estimating the parameters in all the models with Markov chain Monte Carlo (MCMC) technique. In Bayesian inference, parameters are treated as random variables and are given prior distributions. These prior distributions are updated with the likelihood from the collected data to give the posterior distributions of the parameters of interest.

3. Results and Discussion

3.1. Exploring ARI by Various Risk Factors (Bivariate Analysis)

The following risk factors associated with ARI were established in the dataset. They included mother’s age, mother’s education level, type of cooking fuel, wealth index, household size and place of residence. The results that follow show the significance of each of the explanatory variable when logistic regression was carried out with respect to ARI. Table 1 shows the odds ratio estimates of the logistic regression model that was employed to determine the significant risk factors of acute respiratory infection among children less than five years. Place of residence, wealth quantile and household size were found to be significantly associated with ARI because the 95% confidence interval for the odds ratio of these variables did not contain 1 while mother’s age, mother’s education level and type of cooking fuel were found to be insignificant (their 95% confidence interval contained 1) and were dropped from the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR 95% C. I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother’s Age (ref= 15-19 years)</td>
<td>1</td>
</tr>
<tr>
<td>20-24</td>
<td>0.947(0.743,1.21)</td>
</tr>
<tr>
<td>25-29</td>
<td>0.878(0.689,1.116)</td>
</tr>
<tr>
<td>30-34</td>
<td>0.850(0.661,1.097)</td>
</tr>
<tr>
<td>35-39</td>
<td>0.791(0.605,1.03)</td>
</tr>
</tbody>
</table>
From the analysis, wealth quantile was significantly associated with the risk of ARI. 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 (OR: 1.648 (1.324, 2.128)). However, there was a reduction in the risk in children whose households were richest 0.896(0.529, 0.926). A similar study carried out in Zimbabwe by [12] suggested that children from richer communities are less likely to be affected by ARI as compared to those from poor communities. For household size, the risk increased as the number of household members having the highest risk of infection (OR: 1.568 (1.089, 2.248)). This is well support by a study carried out in Kibera, Kenya by [4] who found out that relative risk of ARI cases in children living in a house with between 4-6 persons as 1.05 and those living in households with more than 7 people as 1.24. Therefore, there is a relationship between overcrowding and acquisition of ARI. [13] also noted that houses which had more than three occupants had more ARI cases than those with less. 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. This is also consistent with a similar study carried out in Kenya by which reported that the prevalence of ARI was higher in children living in urban areas as compared to those living in rural areas. The significance of these three variables in relation to ARI was also reported and supported by [10] who carried out an analysis of risk factors for Acute Respiratory Tract Infection (ARTI) in children less than 5 years in Enugu south east of Nigeria.

3.2. Models Comparison

Model estimation was carried out using a Bayesian approach. The spatially structured component was assigned a conditional autoregressive prior and their corresponding precision parameter \( \frac{1}{\delta^2} \) were given a non-informative gamma distribution priors, i.e \( \delta \sim \text{gamma}(0.5, 0.0005) \). All the fixed effect parameters \( \beta \) were given non-informative normally distributed priors, i.e \( \beta \sim \text{Normal}(0.01, 0.01) \).

The models will be implemented using WinBUGS version 1.4 [14]. For each model, 60,000 McMC iterations were ran, with the initial 15,000 iterations discarded to cater for burn period and thereafter keeping every tenth sample value. The 45,000 iterations left were used to assess convergence of the chain and parameter estimation. Convergence of parameters was checked using the trace plots and autocorrelation plots of the MøøMC output, [14]. The BGR statistic was also considered in checking the convergence where values around 1 indicate convergence, with 1.1 considered as acceptable limit by [16].

In this case the value for BGR statistics after running two chains at the 60,000 iteration was 0.998. The models were compared using the Deviance Information Criterion (DIC) as suggested by [17] where the best fitting model is the one with the smallest DIC value. In this analysis, the ARI data for children under five years in Kenya was analyzed using the following models: logistic regression model (M1), the normal unstructured heterogeneity random effects model (M2), the ICAR spatial random effects model (M3) and the convolution model (M4).

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR 95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>40-44</td>
<td>0.851(0.623,1.165)</td>
</tr>
<tr>
<td>45-49</td>
<td>0.905(0.568,1.347)</td>
</tr>
<tr>
<td>Mother’s Education level(ref=no education)</td>
<td>1</td>
</tr>
<tr>
<td>Primary incomplete</td>
<td>0.941(0.815,1.085)</td>
</tr>
<tr>
<td>Secondary complete</td>
<td>0.845(0.656,1.088)</td>
</tr>
<tr>
<td>Higher education</td>
<td>0.878(0.728,1.056)</td>
</tr>
<tr>
<td>type of cooking fuel(ref=electricity/gas)</td>
<td>1</td>
</tr>
<tr>
<td>Paraw</td>
<td>0.926(0.661,1.283)</td>
</tr>
<tr>
<td>Coal</td>
<td>0.783(0.565,1.083)</td>
</tr>
<tr>
<td>Charcoal</td>
<td>0.903(0.758,1.082)</td>
</tr>
<tr>
<td>Wood</td>
<td>1.430(1.374,5.417)</td>
</tr>
<tr>
<td>Animal dung</td>
<td>1.347(0.235,7.959)</td>
</tr>
<tr>
<td>Others</td>
<td>1.099(0.723,1.689)</td>
</tr>
<tr>
<td>household size(ref=1-4 members)</td>
<td>1</td>
</tr>
<tr>
<td>5-9 members</td>
<td>1.413(1.101,1.806)</td>
</tr>
<tr>
<td>10-14 members</td>
<td>1.568(1.089,2.248)</td>
</tr>
<tr>
<td>place of residence(ref=rural)</td>
<td>1</td>
</tr>
<tr>
<td>Urban</td>
<td>1.674(1.321,2.130)</td>
</tr>
<tr>
<td>Wealth quantile(ref=richest)</td>
<td>1</td>
</tr>
<tr>
<td>Poorest</td>
<td>1.648(1.324,2.128)</td>
</tr>
<tr>
<td>Poorer</td>
<td>1.428(1.240,1.632)</td>
</tr>
<tr>
<td>Middle</td>
<td>1.318(1.134,1.592)</td>
</tr>
<tr>
<td>Richer</td>
<td>0.896(0.529,0.926)</td>
</tr>
</tbody>
</table>

Table 2. Model comparison in mapping ARI in Kenya.
From 2 above, comparing the models based on the DIC, the convolution model (DIC = 6914.550) performed better than simple logistic model (DIC = 7234.910), the unstructured random effects model (DIC = 6932.650) and spatial random effects model (DIC = 6935.650). Following these results, the convolution model was selected as the best model for mapping ARI in Kenya.

3.3. Mapping of ARI Across Different Counties in Kenya

Figure 1. Map showing prevalence of ARI across different counties in Kenya.

3.4. Discussion

This study used a Bayesian hierarchical spatial modelling approach to investigate the spatial variation in the prevalence of ARI in Kenya. Bayesian inference is commonly used to produce stabilized risk maps through borrowing information from neighboring regions across the map [18]. Most spatial data sets especially those obtained from geo-demographic and health surveys, not only possess global spatial autocorrelation but they also exhibit significant patterns of spatial instability, which is related to regional differences with the observational space.

The convolution model was the best fitting model and was therefore used to produce county specific map of prevalence of ARI in Kenya. Maps play an important role in understanding disease epidemiology and also in guiding policy makers to develop intervention programs that are most importantly needed by people and allocate the scarce resources appropriately. This study has demonstrated both geographical heterogeneity and the high prevalence rate of ARI in Kenya. The national preference of ARI in Kenya was found to be 47.3%. This rate is very worrying and as projected by [5] if the present trends continue, 4.4 million children younger than five years will still die in 2030 of which majority of these deaths (60%) will occur in sub Saharan Africa.

These results are in line with a study conducted by [19] who found out that the prevalence of ARI for children under five years in Kenya was between 21.7% and 40% in 2011. From Fig 1 above, the results of the study indicate high prevalence levels and clusters of the disease in counties in western part of Kenya which included Bungoma county (shown by red shading) with the highest level (0.840 (0.794, 0.888)) followed by Vihiga county (0.699 (0.628, 0.779)) and Kakamega county (0.697 (0.628, 0.775)) (shaded in plum). The counties which had the lowest prevalence levels were Samburu county (0.184 (0.131, 0.259)), Kirinyaga county (0.235 (0.153, 0.362)), Turkana county (0.289 (0.117, 0.480)), Meru county (0.293 (0.213, 0.436)) and Embu county 0.297 (0.217, 0.386)) (shaded in blue). These results are also similar to the report released by the Kenya Demographic and Health Survey of 2014 which ranked Bungoma county and Vihiga county as the counties with the highest prevalence of ARI and also Turkana and Kirinyaga counties as among the counties with the lowest prevalence of ARI. We can clearly see that there is evidence of spatial correlation as counties which are close together tend to have almost the same rate of ARI prevalence. This study demonstrates the substantial variation in prevalence of ARI among young children in different settings. The difference between the prevalence rates in the different counties could be very much attributed to the difference in socio economic status as well as the climatic conditions of these areas. The result of this study provided useful information on the prevailing epidemiological situation of ARI in all the forty seven counties in Kenya. Considering that the national prevalence of ARI was 47.3%, 18 counties are above this value and it’s important for the county governments of these counties to intervene and provide a solution as to what can be done to reduce the elevated risk levels.

4. Conclusions and Recommendations

There is still much that need to be done to fight against ARI in Kenya. Many counties should consider putting measures in place in order to curb this menace. From his study, it’s clear that socio-economic status affect the risk of death from ARI and should therefore be included in the strategies to reduce mortality from ARI. Disease prevalence is often associated with many socio-economic status factors such as overcrowding, unemployment rates, educational and housing quality among other factors. There is need to intervene against risk factors to prevent clinical pneumonia and significantly reduce the large burden of disease in childhood resulting from ARI. Resources need to be directed where there are much needed in order to avoid wastage of resources. There is need to look into this disease with a lot of seriousness. Parents should also be on the lookout of the clinical signs of ARI in order to seek medical attention early enough. Interventions by the national and county governments would greatly be needed if Kenya will achieve progress in the health sector as a way of achieving vision 2030.
Acknowledgements

I would like to sincerely thank God for the far that I have come with this work. It's by His grace that I have made it to the end. Special thanks to my supervisor Dr. Oscar Owino Ngesa for sharing his knowledge generously in the field of spatial statistics and always replying timely to my many questions even during his resting hours. God bless you abundantly Dr. Ngesa. I would also want to acknowledge the endless support of my family especially my parents Mr. and Mrs. John Kinyua, my sisters Jecinta and Lucy and my close friend Dafton Wilson Wahuga. Finally, I would like to thank my classmates and all the lectures of Jomo Kenyatta University of Agriculture and Technology for the assistance accorded to me in the course of my studies, may God bless you all.

References