LANCASTER UNIVERSITY

Modelling Malnutrition among Under-Five-Year-Old Children in Ghana

Justice Moses Kwaku Aheto
Faculty of Health and Medicine
Medical School

This thesis is submitted for the Degree of

Doctor of Philosophy

January 2016

DECLARATION OF AUTHORSHIP

I, Justice Moses Kwaku Aheto, declare that the work in this thesis titled 'Modelling Malnutrition among Under-Five-Year-Old Children in Ghana' is my own, and I confirm the following:

- The work in this thesis was not submitted for the award of a higher degree elsewhere.
- The work in this thesis was done wholly in my candidature for a research degree at Lancaster University.
- I have clearly indicated any part of this thesis that has been previously published or submitted for publication in scholarly journal.
- Where there are multiple authors in the published/submitted works that constitute the work in this thesis, my responsibilities are clearly stated and these responsibilities approved by my supervisors/co-authors.
- Where I have consulted the published work by others, this is clearly cited accordingly.
- Where I have quoted from the work of others, the source is always provided. Apart from such quotations, the work in this thesis is entirely my own.

Signed: ………………………………………………………………………….

Date: …………………………………………………………………………

“We are guilty of many errors and
many faults, but our worst crime
is abandoning the children,

neglecting the foundation of life.

Many of the things we need can wait.

The child cannot.

Right now is the time his bones are

being formed, his blood is being made

and his senses are being developed.

To him we cannot answer “Tomorrow”.
His name is “Today”. ”

~Gabriela Mistral, 1948

LANCASTER UNIVERSITY

Abstract

Faculty of Health and Medicine

Medical School

Doctor of Philosophy

Justice Moses Kwaku Aheto

Childhood malnutrition is a real-life and a chronic problem and one of the global major public health challenges, especially in developing countries like Ghana. Several attempts from governmental and non-governmental organizations to address the problem have fallen below expectation. It is recognised that the existing studies and nutrition intervention strategies are inadequate and hence not working to expectation. This thesis examines childhood malnutrition in Ghana using appropriate and advanced statistical methods to help improve the understanding of childhood nutrition and to better inform targeted public health nutrition interventions in the country.

In this thesis, we provided solutions to five main problems: (1) investigated the major risk factors for malnutrition; (2) investigated household level variations in nutritional outcomes of children; (3) explored, modelled and illustrated spatial variations in the risk of childhood malnutrition over Ghana; (4) explored, modelled, forecasted and illustrated spatio-temporal variations in the risk of childhood malnutrition over Ghana; (5) jointly modelled weight-for-age Z-score (WAZ) and height-for-age Z-score (HAZ) to improve accuracy and reliability in estimates. To answer the first and the second problems, multilevel models were considered. The results showed strong residual household-level variations in under-fives nutritional outcomes and that child's age, type of birth, child's experience of diarrhoeal episodes, size of child at birth and months of breast feeding, mother's education, current age, BMI and national health insurance status, household toilet facility ownership and wealth status were predictive of under-fives nutrition.

To answer the third problem, spatial models were employed. The study found substantial spatial variation in the predicted risk of under-fives malnutrition over Ghana and also showed that Normalised Difference Vegetation Index (a marker for vegetation

cover), elevation and rural/urban residence status were predictive of under-fives nutritional outcomes. The study considered spatio-temporal models to answer the fourth problem. The results showed substantial spatio-temporal variation in the risk of under-fives chronic malnutrition over Ghana. Our forecasted map of chronic malnutrition showed substantial spatial variation with children from parts of Northern and Western regions being at the highest risk of malnutrition compared to children from other regions of the country. In our forecast maps, the effect of increasing the level of maternal education was shown to reduce the prevalence of malnutrition throughout Ghana.

To answer the fifth problem, multivariate response multilevel models were considered. The study found that the residual household effects for WAZ and HAZ are very strongly correlated and that the correlation was stronger for the residual household effects than the residual child effects. This also suggests that after adjusting for risk factors in our model, it is the same as-yet unidentified factors at household level that influence both WAZ and HAZ. The results also showed that there was more accuracy and reliability in estimates from the multivariate response multilevel model over separate multilevel models and showed that the effect of some important risk factors differed substantially across WAZ and HAZ.

The findings from this thesis are intended to help policymakers responsible for the health and nutrition of children to design efficient public health policies and targeted nutrition interventions amidst scarce public health resources available in Ghana to better understand, target and to reduce childhood malnutrition prevalence closer to the level expected in a healthy, well-fed population of children under-fives.

Acknowledgements

First of all, I would like to thank my supervisors, Dr. Thomas J. Keegan, Dr. Benjamin M. Taylor and Prof. Peter J. Diggle, for their expertise and encouragements throughout my PhD program and for providing me the needed environment to undertake my study successfully here at Lancaster and it is my fervent hope to give same back to my students, if not more.

I would like to thank my internal examiner Dr. Jonathan Read and my external examiner Dr. Marko Kerac for their careful reading, constructive criticisms and time. Also, I would like to thank Dr. Christopher Jewell for being the independent chair of my viva.

Being part of CHICAS here at Lancaster Medical School is one of the best things that ever happened in my life and I benefitted heavily in the knowledge base among members which I gratefully acknowledge. CHICAS seminars were very informative and educating and provided the needed platform for members to interact and exchange ideas with similar minded persons on our research areas. In many similar places, such opportunities are rare.

I am grateful to all my families back in Ghana who managed to cope with my absence while I am away on this all important PhD program abroad.

I can’t end my acknowledgements successfully without acknowledging the financial support for my PhD from the Faculty of Health and Medicine here at Lancaster University as well as the Ghana Statistical Service and the DHS MEASURE Program for providing the data for my PhD Studies. I am also grateful to William Ritchie Travel Fund here at Lancaster University for sponsoring the presentation of my paper titled “Analysing Malnutrition Prevalence and its Determinants among Under-Five Children in Ghana: Multilevel Methods” at ‘Nutrition and Nurture in Infancy and Childhood: Bio-Cultural Perspectives Conference’ here in the United Kingdom.

Remarks

This study has provided me a great opportunity to further appreciate more than ever the need for statistical modelling in areas related to epidemiology and public health. I have learnt through this work that statistical methods will continue to be indispensable in epidemiology and public health by guiding and supporting policy decisions and intervention strategies. The study has also equipped me with appropriate and advanced statistical modelling techniques and it is my hope to put these skills to the best and maximum use in order to contribute my quota to the area of statistical modelling, public health and epidemiology as well as other related disciplines.

This study has also reminded me of the fact that one should make the best and the maximum use of his or her expertise to contribute meaningfully to solving some challenges facing the society and that such effort will not go unrecognised by the society. This was manifested in the prestigious 'TakeAIM Prize' that I have won in 2015 based on the work in this thesis which was recognised for its potential impact in the study of child malnutrition by Smith Institute for Industrial Mathematics and System Engineering, London, UK.

Contents

DECLARATION OF AUTHORSHIP i
Abstract iii
Acknowledgements v
Remarks vi
List of Tables xi
List of Figures xii
Chapter 1 General Introduction 1
1.1 Background 2
1.2 Ghana's history, geography, population and economy 3
1.3 Global burden of under-five child malnutrition 9
1.4 Under-five malnutrition trends in Ghana 10
1.5 Major risk factors for under-five child malnutrition 12
1.6 The need for this study 16
1.7 Study objectives, research questions, contributions and organization 20
1.8 Role in the published/submitted/non submitted works 24
Reference 25
Chapter 2 The Ghana Demographic and Health Survey Data 28
2.1 Introduction 29
2.2 Survey designs 29
2.3 Data sets used in this thesis 31
2.3.1 Data quality and limitations 32
2.3.2 Variables used in the analysis 35
2.3.3 Process of selecting explanatory variables used in this thesis 36
2.3.4 Comparing the GDHS to other DHS countries 36
Reference 38
Chapter 3 General Overview of Statistical Methods 39
3.1 Multilevel and multivariate response multilevel modelling 40
3.1.1 Multilevel modelling 40
3.1.2 Multivariate response multilevel modelling 42

3.2 Spatial and spatio-temporal modelling 44
3.2.1 Spatial modelling 44
3.2.2 Spatio-temporal modelling 47
3.2.3 The filtering recursion 47
3.2.4 The Kalman Filter 48
3.6 Conclusion 50
Reference 51
Chapter 4 Paper 1: Childhood Malnutrition and its Determinants among Under-Five Children in Ghana 52
4.1 Introduction 55
4.2 Methods 56
4.2.1 Study population 56
4.2.2 Outcome variables 57
4.2.3 Explanatory variables 58
4.2.4 Statistical analysis 58
4.3 Results 59
4.3.1 Sample characteristics 59
4.3.2 Risks factors associated with WAZ, HAZ and WHZ for multilevel models 61
4.3.3 Probabilities of malnutrition for significant risk factors 63
4.4 Discussion 64
4.5 Conclusion 72
4.6 Appendix 73
4.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis. 73
4.6.2 Online Supporting Information Appendix S2: Supplementary material related to the random intercept multilevel regression model and computation of probabilities of malnutrition 73
4.6.3 Online Supporting Information Appendix S3: Supplementary material relating to the logistic regression analysis of the association between self-reported birth size (large/average versus small) and actual birthweight. 76
4.7 Annex 1: STROBE checklist 77
4.8 Annex 2: Assessing severity of malnutrition by prevalence ranges 79

Chapter 5 Paper 2: Modelling Spatial Variations in the Risk of Childhood Malnutrition among Under-Five Children in Ghana 83

5.1 Introduction 86

5.2 Methods 87

5.2.1 Study population, design and sample 87

5.2.2 Outcome variable 88

5.2.3 Explanatory variables 89

5.2.4 Statistical analysis 89

5.3 Results 90

5.3.1 Sample characteristics 90

5.3.2 Risks factors associated with WAZ, HAZ and WHZ for Spatial mixed models 92

5.3.3 Spatial variation in risk of malnutrition 93

5.4 Discussion 97

5.5 Conclusion 102

5.6 Appendix 103

5.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis. 103

5.6.2 Supporting Information Appendix S2: Supplementary material related to the spatial model and constructing predictive maps for probabilities of malnutrition 103

5.7 Annex 1: STROBE checklist 107

Chapter 6 Paper 3: Modelling and Forecasting Spatio-temporal Variations in the Risk of Chronic Malnutrition among Under-Five Children in Ghana 111

6.1 Introduction 114

6.2 Methods 115

6.2.1 Study population and design 115

6.2.2 Geographical data 116

6.2.3 Outcome variable 117

6.2.4 Explanatory variables 117

6.2.5 Numbers of eligible participants 118

6.2.6 Statistical analysis 118

6.3 Results 124

6.4 Discussion 128

6.5 Conclusion 131

6.6 Appendix 133

6.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis. 133

6.6.2 Supporting Information Appendix S2: Supplementary material related to numbers of children included/excluded for each of the four data collection rounds used in the study. 133

6.7 Annex 1: STROBE checklist 134

Chapter 7 Paper 4: Joint Modelling of Childhood Nutritional Outcomes among Under-Five Children in Ghana 139

7.1 Introduction 142

7.2 Methods 144

7.2.1 Study population 144

7.2.2 Outcome variables 145

7.2.3 Explanatory variables 145

7.2.4 Statistical analysis 145

7.3 Results 148

7.3.1 Correlates for weight-for-age and height-for-age Z scores for multivariate response multilevel models 148

7.4 Discussion 151

7.5 Conclusion 157

7.6 Appendix 158

7.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis. 158

7.6.2 Supporting Information Appendix S2: Information relating to our multivariate normality diagnostics for our multivariate response multilevel model for WAZ and HAZ, and WAZ, HAZ and WHZ scores. 158

7.7 Annex 1: STROBE checklist 160

Chapter 8 General discussion and future studies 164

List of Tables

TABLE 1.1 UNDER-FIVE MALNUTRITION TRENDS FOR GHANA 11

TABLE 1.2 MALNUTRITION PREVALENCE FOR GHANA AND HER NEIGHBOURING COUNTRIES 11

TABLE 1.3 STUDY OBJECTIVES AND RESEARCH QUESTIONS 20

TABLE 2.1 TIME OF INTERVIEW PERIOD FOR DATA SETS USED IN THIS THESIS 31

TABLE 2.2 DESCRIPTION OF VARIABLES USED IN THIS THESIS 35

TABLE 2.3 SUMMARY RESULTS OF SOME SELECTED DHS 2008-2011 37

TABLE 4.1 PERCENTAGE DISTRIBUTION OF CHILDREN UNDER-FIVE BY SELECTED BACKGROUND CHARACTERISTICS (N=2083) 60

TABLE 4.2 THE EFFECT ESTIMATE (Β) FOR THE ASSOCIATIONS BETWEEN RISK FACTORS AND CHILD NUTRITIONAL STATUS (WEIGHT-FOR-AGE OR WAZ (A MEASURE FROM WHICH UNDERWEIGHT IS DERIVED), HEIGHT-FOR-AGE OR HAZ (A MEASURE FROM WHICH STUNTING IS DERIVED) AND WEIGHT-FOR-HEIGHT OR WHZ (A MEASURE FROM WHICH WASTING IS DERIVED)) AMONG UNDER-FIVE CHILDREN IN GHANA USING RANDOM INTERCEPT MULTILEVEL REGRESSION MODELS (N=2083). 62

TABLE 4.3 VARIANCE ESTIMATES WITH 95% CONFIDENCE INTERVALS FOR THE RANDOM INTERCEPT MULTILEVEL REGRESSION MODELS OF WEIGHT-FOR-AGE OR WAZ (A MEASURE FROM WHICH UNDERWEIGHT IS DERIVED) , HEIGHT-FOR-AGE OR HAZ (A MEASURE FROM WHICH STUNTING IS DERIVED), AND WEIGHT-FOR-HEIGHT OR WHZ (A MEASURE FROM WHICH WASTING IS DERIVED) (N=2083) 64

TABLE 4.4 DESCRIPTION OF EXPLANATORY VARIABLES USED IN THIS ANALYSIS 73

TABLE 4.5 THE EFFECT ESTIMATE (Β) FOR THE ASSOCIATIONS BETWEEN SIZE OF CHILD AT BIRTH (OUTCOME VARIABLE) AND ACTUAL BIRTHWEIGHT FROM LOGISTIC REGRESSION MODEL (N=830*). 76

TABLE 5.1 PERCENTAGE DISTRIBUTION OF MALNOURISHED/NOT MALNOURISHED CHILDREN BY SELECTED BACKGROUND CHARACTERISTICS (N=2051). 91

TABLE 5.2 THE EFFECT ESTIMATE (Β) FOR THE ASSOCIATIONS BETWEEN RISK FACTORS AND CHILD NUTRITIONAL STATUS (WEIGHT-FOR-AGE OR WAZ, HEIGHT-FOR-AGE OR HAZ, AND WEIGHT-FOR-HEIGHT OR WHZ) AND SPATIAL EFFECT ESTIMATES FOR SPATIAL MIXED MODELS (N=2051). 93

TABLE 5.3 DESCRIPTION OF EXPLANATORY VARIABLES USED IN THIS ANALYSIS 103

TABLE 6.1 THE EFFECT ESTIMATE (Β) FOR THE ASSOCIATIONS BETWEEN RISK FACTORS AND CHILD NUTRITIONAL STATUS (HEIGHT-FOR-AGE Z-SCORE OR HAZ SCORE) FOR SPATIOTEMPORAL MODEL (N=10,036). 124

TABLE 6.2 PARAMETERS OF THE FITTED SPATIOTEMPORAL MODEL FOR HEIGHT-FOR-AGE Z-SCORE (OR HAZ) 125

TABLE 6.3 DESCRIPTION OF EXPLANATORY VARIABLES USED IN THE ANALYSIS 133

TABLE 7.1 PARAMETER ESTIMATES FOR MULTIVARIATE RESPONSE MULTILEVEL MODELS OF WEIGHT-FOR-AGE (WAZ) VERSUS HEIGHT-FOR-AGE (HAZ) Z-SCORES. 149

TABLE 7.2 VARIANCE AND RESIDUAL CORRELATION ESTIMATES FOR THE MULTIVARIATE RESPONSE MULTILEVEL MODELS OF WEIGHT-FOR-AGE VERSUS HEIGHT-FOR-AGE Z-SCORES. 150

TABLE 7.3 TEST FOR EQUALITY OF RISK FACTOR EFFECT ACROSS WAZ AND HAZ SCORES. 151

TABLE 7.4 DESCRIPTION OF EXPLANATORY VARIABLES USED IN THIS ANALYSIS 158

List of Figures

FIGURE 1.1: MAP OF GHANA 5
FIGURE 1.2: PERCENTAGE SHARE OF POPULATION BY REGION 6
FIGURE 1.3: POPULATION BY TYPE OF LOCALITY (URBAN/RURAL) 8
FIGURE 1.4: ESTIMATED PREVALENCE OF UNDERWEIGHT, STUNTING, AND WASTING DURING 2011 BY REGION. 10
FIGURE 1.5: FRAMEWORK OF THE ASSOCIATION BETWEEN POVERTY, FOOD INSECURITY, AND OTHER UNDERLYING AND IMMEDIATE CAUSES TO UNDER-FIVE CHILD MALNUTRITION AND ITS ASSOCIATED SHORT AND LONG TERM CONSEQUENCES. 12
FIGURE 2.1 THE HIERARCHICAL STRUCTURE OF THE DATA-SETS USED IN THIS THESIS 32
FIGURE 3.1: STATE SPACE MODEL DEPENDENCE STRUCTURE 47
FIGURE 4.1 PLOTS SHOWING PROBABILITIES OF MALNUTRITION FOR CONTINUOUS RISK FACTORS, BY MALNUTRITION MEASURE: A) UNDERWEIGHT (WAZ), B) STUNTING (HAZ), C) WASTING (WHZ)... 65
FIGURE 4.2 BOX PLOTS SHOWING PROBABILITIES OF MALNUTRITION FOR CATEGORICAL RISK FACTORS, BY MALNUTRITION MEASURE: A) UNDERWEIGHT (WAZ), B) STUNTING (HAZ), C) WASTING (WHZ)... 67
FIGURE 4.3 PROBABILITY PLOT OF BEING BORN LARGE/AVERAGE VERSUS SMALL IN SIZE AT BIRTH AGAINST ACTUAL BIRTHWEIGHT 76
FIGURE 5.1 GHANA MAP SHOWING DISTRIBUTION OF CENTROIDS (.) OF THE SAMPLED CLUSTERS DURING THE GHANA DEMOGRAPHIC AND HEALTH SURVEY CONDUCTED IN 2008 94
FIGURE 5.2 SPATIAL VARIATIONS IN PROBABILITIES OF UNDERWEIGHT IN CHILDREN <5 IN GHANA..... 95
FIGURE 5.3 SPATIAL VARIATIONS IN PROBABILITIES OF STUNTING IN CHILDREN <5 IN GHANA 96
FIGURE 5.4 SPATIAL VARIATIONS IN PROBABILITIES OF WASTING IN CHILDREN <5 IN GHANA 97
FIGURE 6.1 PROBABILITIES OF STUNTING IN GHANA IN 1993 (TOP LEFT), 1998 (TOP RIGHT), 2003 (BOTTOM LEFT) AND 2008 (BOTTOM RIGHT) IN CHILDREN <5 YEARS 126
FIGURE 6.2 OVERALL PROBABILITIES OF STUNTING FORECAST IN GHANA IN 2013 IN CHILDREN <5 YEARS 127
FIGURE 6.3 EFFECT OF A MINIMUM OF 6 (LEFT) AND 12 (RIGHT) YEARS OF MOTHERS' EDUCATION ON PROBABILITIES OF STUNTING FORECAST IN GHANA IN 2013 IN CHILDREN <5 YEARS..... 128
FIGURE 6.4 DATA-SETS USED IN THIS STUDY 133
FIGURE 7.1 MULTIVARIATE NORMALITY PLOT FOR WAZ AND HAZ SCORES. 158
FIGURE 7.2 MULTIVARIATE NORMALITY PLOT FOR WAZ, HAZ AND WHZ SCORES 159

I dedicate this work to my son, wife and mother and to my late grandmother who passed on in September, 2015 during the last year of this PhD work. Also, I dedicate this work to my late father, to my siblings and to the world children.

Chapter 1 General Introduction

1.1 Background

Childhood malnutrition is a chronic problem and one of the major public health challenges, especially in developing countries like Ghana. Several attempts from governmental and non-governmental organizations to address the problem have fallen below expectation. In 2008, estimates indicate that 28%, 14% and 9% of children under the age of 5 years are stunted (chronic malnutrition), underweight and wasted (acute malnutrition) respectively in Ghana (Ghana Statistical Service et al., 2009). This could pose a major threat to the growth, survival and development of Ghanaian children.

Childhood is a period associated with an active growth, covering the major transformations from birth to adulthood and as a result requires healthy and adequate nutrition to maintain optimum growth and development in these children. The contribution of malnutrition to morbidity and mortality has been shown to be synergistic rather than additive (Pelletier DL et al., 1993). Thus, mortality increases exponentially with deteriorating nutritional outcomes.

The high levels of under-fives malnutrition in Ghana led the Ghanaian Government to develop the 2014 National Nutrition Policy in an attempt to: (1) reposition nutrition as a cross-cutting issue; (2) provide the framework for nutrition-specific and nutrition-sensitive services and interventions; (3) facilitate integration and mainstreaming of nutrition into national development efforts; (4) strengthen sectorial capacity for the effective delivery of these interventions; and (5) guide the implementation of high-impact interventions. In 2011, the Government of Ghana also joined the Scaling Up Nutrition (SUN) Movement which is a global movement uniting national leaders, civil society, the United Nations, donors, businesses and researchers in an attempt to address malnutrition. The SUN movement focus on targeting resources to mothers and children with emphasis on the first 1000 days of life starting from pregnancy.

Addressing the problem of childhood malnutrition in Ghana is expected to contribute to significant reduction in under-fives mortality rate in the country as well as preparing the children for better developmental paths and increasing national productivity and sound health in adult life (The World Bank, 2006). The under-five mortality rate in Ghana is 80 per 1,000 live births in the year 2008 (Ghana Statistical Service et al., 2009).

The few studies (Antwi, 2008; Ghana Statistical Service et al., 2009) conducted in Ghana to examine childhood nutrition have not always used appropriate statistical models to examine the risk factors associated with children's nutritional outcomes. In addition, they did not examine the impact of place and time on under-fives nutrition. Furthermore, since malnutrition prevalence persist in Ghana, continued examination of the malnutrition prevalence and its associated risk factors on the outcome using appropriate and advanced statistical methods is warranted if an appreciable level of progress is to be made in reducing the prevalence closer to those expected in a healthy, well-fed under-five children population.

In this chapter, we present the history of Ghana, her geography, population and socioeconomic background, the global under-five child malnutrition, under-five malnutrition trends in Ghana, major risk factors for under-five child malnutrition, the need for the study, study objectives and research questions, contributions and organisation of the thesis.

1.2 Ghana's history, geography, population and economy

Ghana gained independence in 1957 from the British. Ghana is located on the West African Coast and approximately covers a total land area of 238,537 square kilometres. The country is bordered on the north by Burkina Faso, east by Togo and Cote d'Ivoire on the West. The Gulf of Guinea which forms a coastline extending approximately 560 kilometres is located on the Southern part of Ghana.

Ghana is a lowland country but characterised with few range of hills on the eastern border and Mountain Afadjato, which is the highest point above sea level, at 884 metres. The country can be divided into three ecological zones: northern savannah, which is drained by Black and White Volta Rivers, the sandy coastline backed by a coastal plain, which is crossed by several rivers and streams and the middle belt and western part of the country, which are heavily forested and have many streams and rivers. The Volta River is one of the largest man-made lakes in the world which was created by the hydroelectric dam located in the east of the country.

In Ghana, temperatures and rainfalls vary according to distance from the coast and elevation as a result of her tropical climate. The average annual temperature is about 26^{o}C (79^{o}F) with two distinct rainy seasons: April to June and September to November. However, in the north, there is a shift in rainfall pattern and the rainy season begins in March and lasts until September. The annual rainfall ranges from 1,015 millimetres in the north to 2,030 millimetres in the southwest. A dry desert wind known as harmattan blows from the northeast between December and March. It lowers the amount of humidity and creates very warm days and cool nights in the north. However, in the south, the effects of the harmattan are mainly felt in January.

Administratively, there are ten regions in Ghana, each having their own regional capitals with the regions further subdivided into districts according to population size. There are presently 216 districts in the country as at 2015. Each of the districts also has their own district capitals. The reason for the demarcation of Ghana into regions and districts is to promote decentralization in the governance of the country through delegation of power to regional and local authorities by the president of the Republic of Ghana.

Figure 1.1: Map of Ghana
Source: http://www.nationsonline.org/oneworld/map/ghana_map.htm (accessed on 07/05/2015).

The total population of Ghana as reported in the 2010 Population and Housing census is 24,658,823 (Ghana Statistical Service, 2012). This represents an increase of 30.4% over that of the 2000 Population and Housing census population figure of 18,912,079.

The report indicates an annual average intercensal growth rate of 2.5%, suggesting that Ghana's population could double by 28 years to come at this rate (Ghana Statistical Service, 2012). Presented in Figure 1.2 is the percentage share of population by the ten regions in Ghana.

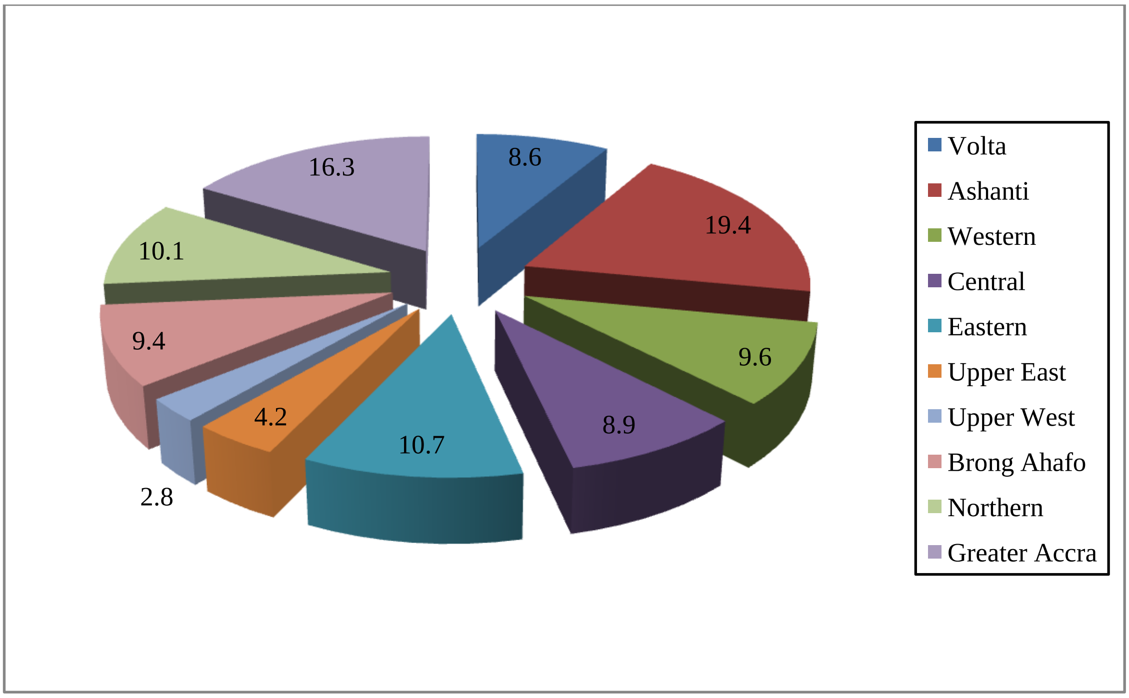


Figure 1.2: Percentage share of population by region
Source: Ghana Statistical Service 2012.Ghana 2010 Population and Housing Census.

Figure 1.2 shows that Ashanti region has the highest population representing 19.4% (4,780,280) of the total population of Ghana followed by the Greater Accra region representing 16.3% (4,010,054). The least populous regions are Upper West and Upper East representing 2.8% (702,110) and 4.2% (1,046,545) respectively.

The census reveals a male and female population of 12,024,845 and 12,633,978 respectively given a sex ratio (number of males to 100 females) of 95.2 compared to 97.9 in 2000 census. Similar trend exist across all the regions except in the Western region where the number of males and females is approximately equal (100:100).

The report also reveals a rise in the population density from 79 persons per square kilometre in 2000 to 103 persons per square kilometre in 2010. This rise implies more pressure is mounted on the country's infrastructures, existing social amenities, natural resources and other resources in the country. Greater Accra and Central regions experienced the highest population growth rates (3.1%), followed by Northern Region (2.9%), Ashanti (2.7%) and Volta (2.5%). The lowest growth rates were recorded for the Upper East (1.2%) and Upper West (1.9%) regions.

Regional statistics indicates that Greater Accra and Central are the most densely populated regions. Greater Accra has a density of approximately 1,236 persons per square kilometre in 2010 as compared to 895.5 persons per square kilometre in 2000. Central region followed with a population density of 224 persons per square kilometre. The Northern region still remains the most sparsely populated region in Ghana with a population density of 35 persons per square kilometre.

The Ghana Statistical Service (GSS) reported that there are 5,467,136 households in the country in 2010 which indicates an increase of 47.7 percent over the number of households recorded in 2000 which gives an average household size of 4.4, indicating an average decline of 0.7 persons from the average household size of 5.1 recorded in 2000.

Ghana has a youthful population as indicated in the results from the 2010 population and housing census report. According to the GSS, the youthful population consists of a large proportion of children under 15 years (38.3%) with a small proportion of elderly persons aged 65 years and older (4.6%), and further concluded that the age structure of the country's population is basically shaped by the effects of high fertility and decreasing mortality rate in the country (Ghana Statistical Service, 2012).

As indicated in Figure 1.3, the proportion of Ghanaians living in urban and rural areas is 50.9% and 49.1 % respectively. The level of urbanization varies from one region to another. Greater Accra recorded the highest level of urban population (90.5%) while the

Upper West recorded the lowest proportion of urban population (16.3%). According to the GSS, the relatively high urban population in the Greater Accra and Ashanti may partly be linked to the concentration of industries and commercial activities in these regions. The remaining 8 regions are predominantly rural and are below the level of national average for urbanization.

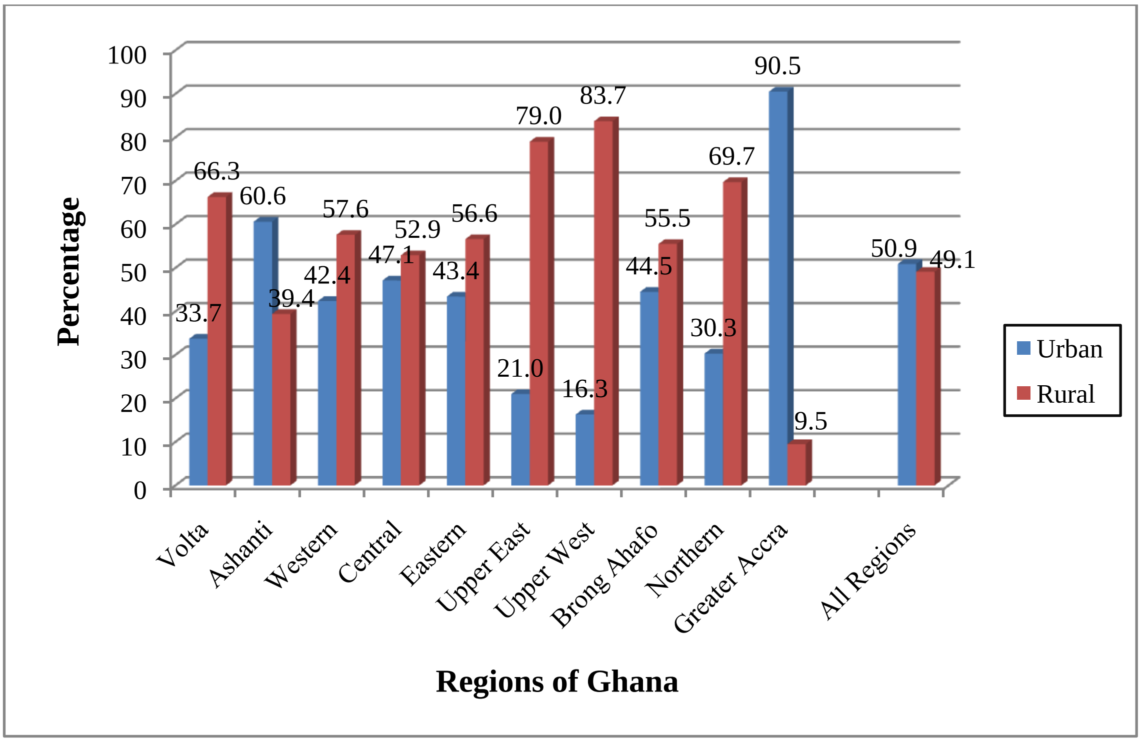


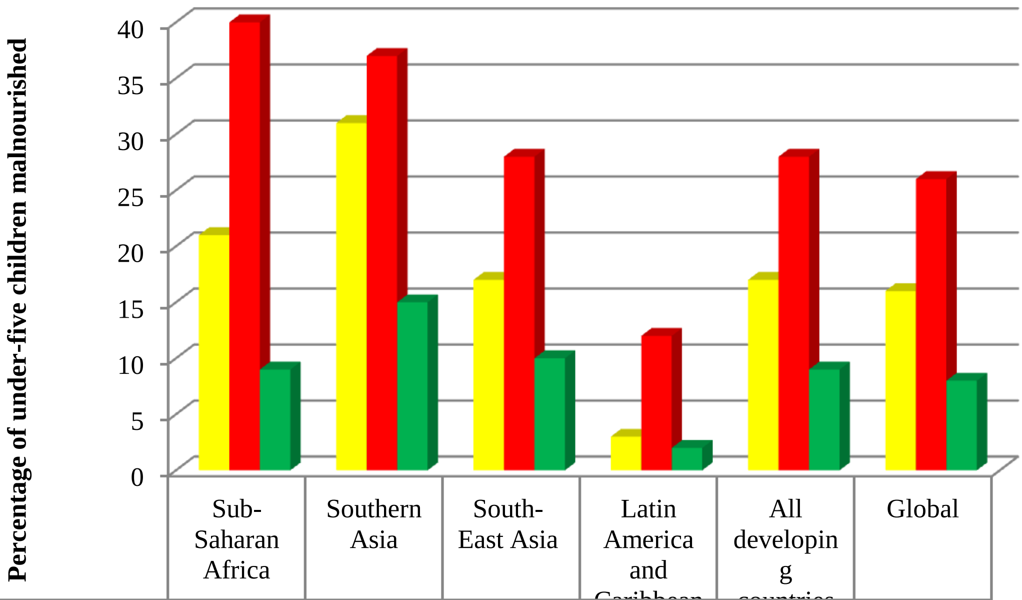
Figure 1.3: Population by type of locality (Urban/Rural)
Source: Ghana Statistical Service 2012. Ghana Population and Housing Census. 2010.

Ghana's gross domestic products (GDP) in 2012 was estimated at 73,109 million Ghanaian Cedi ($22,091,315,646; using the Bank of Ghana exchange rate of $1= GHS3.3094 on 15/07/2015) (Ministry of Food and Agriculture (MoFA), 2013). The agriculture sector remains the backbone of Ghana's economy and used to be the sector that serves as the main employment sector to those in the labour force (Food and Agriculture Organization (FAO), 2009) but the service sector has recently overtaken the agricultural sector. The workers in the agriculture sector are normally poorly paid and their operations are mostly informal in nature with approximately 90% of farm holdings less than 2 hectares in size which suggests that the sector mainly operates on smallholdings. In Ghana like other Sub-Saharan African countries, agricultural production varies according to the amount and distribution of rainfall as well as the soil type available at a given geographical location. Recently, the sector employs 42% of the workforce and contributes 22.7% to gross domestic product (GDP) while the service sector employs 43% of the workforce and contributes 50% to GDP in the year 2012. The industry sector on the other hand employs 15% of the workforce and contributes a total of 27.3% to GDP in same period (Ministry of Food and Agriculture (MoFA), 2013).

1.3 Global burden of under-five child malnutrition

Presented in Figure 1.4 is the global burden of under-fives malnutrition. The prevalence of underweight is highest in Southern Asia representing 31% while stunting on the other hand is more prevalent in Sub-Saharan Africa with the rate of 40%. Southern Asia recorded the highest prevalence of wasting representing 15%. Latin America and the Caribbean recorded the lowest prevalence of underweight, stunting and wasting representing 3%, 12% and 2% respectively. Developing countries recorded underweight, stunting and wasting prevalence of 17%, 28% and 9% respectively while the global burden of underweight, stunting and wasting is estimated at 16%, 26% and 8% respectively.

The global estimated number of under-five-year-old children population in 2014 was 667 million. The 2014 under-five-year-old children population globally showed that 47%, 26% and 15% lived in lower-middle income, low income and upper-middle income countries, respectively. The lower-middle income countries contributed 66% of all stunted children globally whilst low income countries accounted for about 24% of all stunted children. The upper-middle income countries accounted for about 8% of all stunted children globally (UNICEF et al., 2015).



	Sub-Saharan Africa	Southern Asia	South-East Asia	Latin America and Caribbean	All developing countries	Global
Underweight (%)	21	31	17	3	17	16
Stunting (%)	40	37	28	12	28	26
Wasting (%)	9	15	10	2	9	8

Figure 1.4: Estimated prevalence of underweight, stunting, and wasting during 2011 by region. Source: WHO, 2012. http://www.who.int/nutgrowthdb/jme_infosheet.pdf

Globally, under-five child malnutrition rate is decreasing but the progress is slow. However, the number of stunted children in Africa is increasing. Also, none of the sub-regions in Africa has an acceptable level of wasting (UNICEF et al., 2015). There are 34 countries that account for 90% of the global malnutrition burden and some of these countries include Ghana, Nigeria, Bangladesh, Burkina Faso, Cameroon, Angola, Kenya, Uganda, Zambia, Malawi and Iraq among others.

1.4 Under-five malnutrition trends in Ghana

Though malnutrition rates for underweight, stunting and wasting were high in Ghana since 1988, the prevalence of underweight appears to be decreasing appreciably over the years in the country and recorded 14% in 2008 over that of 23% in 1988. Stunting prevalence on the other hand appears to be decreasing from 34% in 1988 to 31% in

1998 but rather peaked up to 35% in 2003 before decreasing to 28% in 2008. Generally, wasting appears not to be decreasing over the years. It increased from 9% in 1988 to 14% in 1993 before decreasing to 10% and 8% respectively in 1998 and 2003 and then increased to 9% in 2008 (see Table 1.1) (Ghana Statistical Service et al., 2009).

It is therefore clear that malnutrition prevalence among children under-fives in Ghana is still high in the country and hence the need to apply appropriate and advanced statistical methods to identify its drivers to inform policy and public health nutrition interventions in the country to reduce the prevalence.

Table 1.1 Under-five malnutrition trends for Ghana

	Malnutrition prevalence (%) by survey years				
Malnutrition indicator	1988	1993	1998	2003	2008
Underweight	23	23	20	18	14
Stunting	34	33	31	35	28
Wasting	9	14	10	8	9

Source: Ghana Statistical Service et al.,2009. Ghana Demographic and Health Survey, 2008

For the purpose of comparison, we have presented malnutrition prevalence for Ghana and her neighbours such as Burkina-Faso, Cote D'Ivoire and Togo for specific years relatively close in time for the study years. Compared to her neighbours, Ghana has the lowest underweight prevalence and has the second lowest prevalence for stunting and wasting after Togo (see Table 1.2). For all the three malnutrition measures, Cote D'Ivoire has the highest prevalence.

Table 1.2 Malnutrition prevalence for Ghana and her neighbouring countries

Country	Year	Underweight(%)	Stunted (%)	Wasted (%)
Ghana	2008	14.3	28.6	8.7
Burkina Faso	2009	26.0	35.1	11.3
Cote D'Ivoire	2007	29.4	39.0	14.0
Togo	2008	20.5	26.9	6.0

Source: WHO Global Database on Child Growth and Malnutrition, 2014
http://www.who.int/nutgrowthdb/database/countries/en/

1.5 Major risk factors for under-five child malnutrition

To investigate the major risk factors for childhood nutritional outcomes, this thesis was guided by UNICEF's conceptual framework to analyse malnutrition in mothers and children which recognises the basic, underlying and immediate causes of malnutrition (UNICEF, 1990). We adapted this validated framework developed by UNICEF to analyse major risk factors for malnutrition in children (see Figure 1.5). The literature review on risk factors for under-five child malnutrition was therefore based on factors described in UNICEF's conceptual framework.

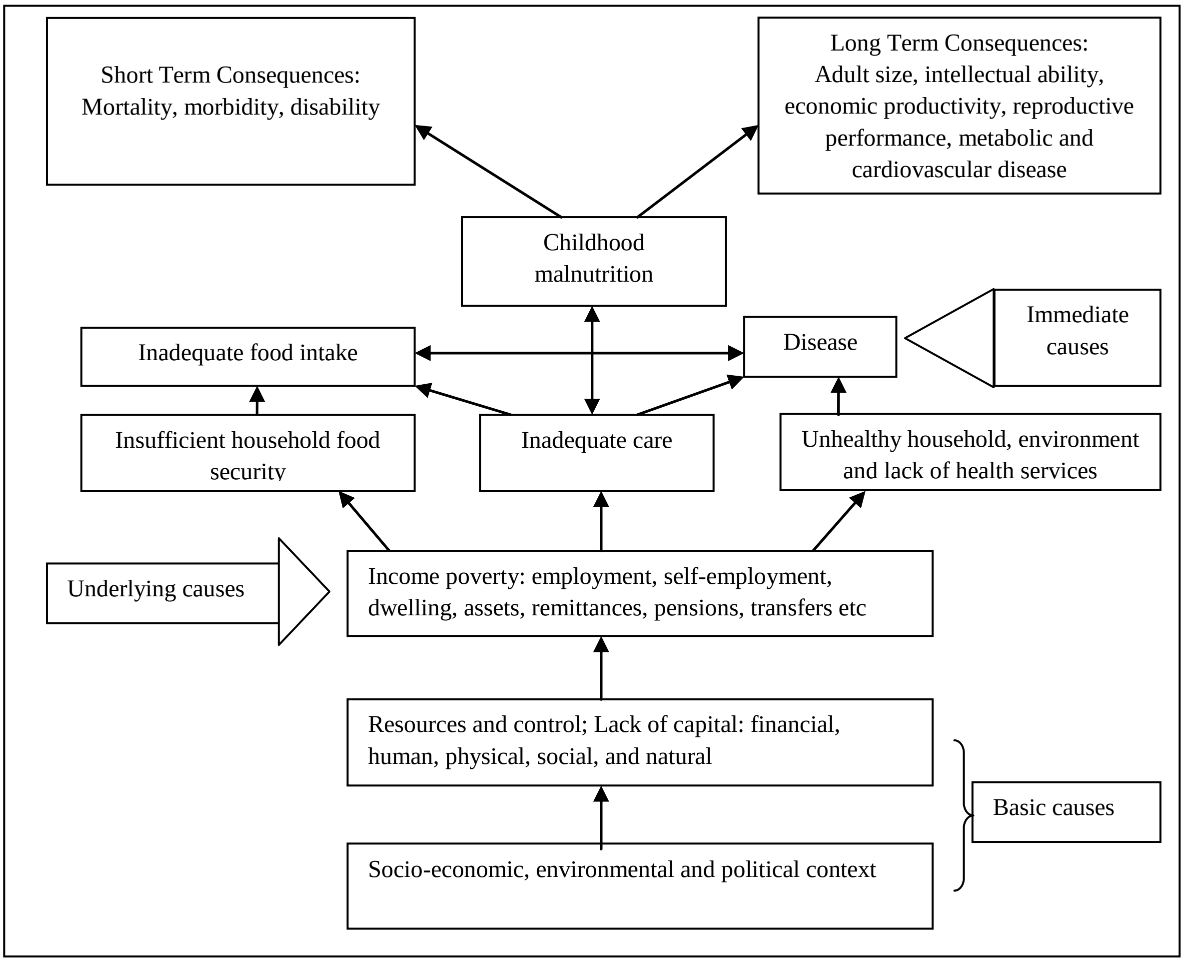


Figure 1.5: Framework of the association between poverty, food insecurity, and other underlying and immediate causes to under-five child malnutrition and its associated short and long term consequences. Note: Figure 1.5 is adapted from UNICEF's Conceptual Framework for the analysis of food and nutrition security (UNICEF 1990, page 22).

In a quest to improve optimum fetal and child nutrition, growth and development to an appreciable level globally, a new conceptual framework was developed and presented in Lancet 2013 that shows the pathways to optimum fetal and child growth and development, outlining the behavioural, health determinants of optimum nutrition, development and growth, and shows how these are influenced by underlying factors such as care-giving resources, environmental conditions and food security, and how these underlying factors are in turn affected by global and national policies, socioeconomic conditions, governance, resources, and capacity (Black et al., 2013).

In the 2013 Lancet nutrition series, differentiation between nutrition-specific and nutrition-sensitive interventions and programmes was highlighted (Bhutta et.al. 2013; Rue et.al. 2013). Nutrition-specific risk factors are the immediate determinants of fetal and child nutrition such as dietary intake, care-giving and parental practices, illness or disease whilst nutrition-sensitive risk factors are the underlying determinants of fetal and child nutrition and development such as safe and hygienic environment, food security, access to health services, care-giving resources at maternal, household and community levels while incorporating specific nutrition goals and actions (Bhutta et.al. 2013; Rue et.al. 2013). The main focus is on how these determinants (nutrition specific and nutrition sensitive) could be modified to promote growth and development. The modification include nutrition-specific interventions that are aimed at addressing the immediate causes of retarded growth and development and the potential effects of nutrition-sensitive interventions that address the underlying determinants of malnutrition while incorporating specific-nutrition goals and actions (Black et al., 2013). This differentiation is important as it will help determine which interventions and programmes will help address nutrition-specific (e.g. breastfeeding and complementary feeding, adolescent health, maternal dietary supplementation, disease prevention and management, etc.) and nutrition-sensitive (e.g. food security, child protection, water and sanitation, health and family planning services, etc.) risk factors influencing optimal fetal and child nutrition, growth and development.

The analysis in this thesis also takes into account the fact that factors that drive a child's nutritional outcome are at four different levels; child level, maternal/household level, community level, and national level. Child level factors such as child's illnesses, dietary intake and biological factors such as age and birthweight of the child play a direct and a crucial role in influencing a child's nutritional outcome whereas maternal/household level factors such as access to resources (income, food, safe water, e.t.c.) and health services (maternal/household health seeking behaviour, child care and feeding practices) are important at influencing a child's dietary intake and their overall health outcomes and therefore play an indirect role in influencing a child's nutritional outcome. Beliefs/customs, vegetation, distance to health facilities and access to health facilities are some community level variables that influence the household level factors. National level factors on the other hand are necessary to influence the implementation of programmes and policies that may be implemented to affect changes to the nutritional outcome of children.

Many studies into childhood nutritional outcomes and its determinants in developing countries established that environmental conditions, biological factors, socio-economic circumstances coupled with feeding practices are very important determinants of under-five child nutritional status in such countries (Adekanmbi et al., 2013; Alom et al., 2012; Babatunde et al., 2011; Bomela, 2009; Engebretsen et al., 2008; Kabubo-Mariara et al., 2009; Kandala et al., 2011; Madise and Mpoma, 1997; Mbago and Namfua, 1992; Saha et al., 2009; Vella et al., 1992; Wamani et al., 2006).

Factors that affect childhood nutrition in Africa are generally understood to be: child's age and sex, type of birth (multiple/singleton), maternal education and body mass index, maternal health-seeking behaviour, household socioeconomic status, child birthweight, source of drinking water, toilet facilities, public health services, disease episodes (e.g. fever, diarrhoea, malaria), geographic differences/place of residence (urban/rural), vegetation and birth intervals (Adekanmbi et al., 2013; Babatunde et al., 2011; Engebretsen et al., 2008; Kabubo-Mariara et al., 2009; Kinyoki et al., 2015; Madise and Mpoma, 1997; Mbago and Namfua, 1992; Wamani et al., 2006).

In Asia, most of the factors found to affect nutrition among children under-five are similar to those found in Africa: child's age, maternal education and body mass index, household socioeconomic status, currently breast-feeding, toilet facilities, number of people in household, number of children under-five in household, months of breastfeeding, paternal education, place of delivery, elevation, father's occupation, household food security, knowledge of oral rehydration therapy and child birthweight (Alom et al., 2012; Bomela, 2009; Dang et al., 2004; Debnath and Bhattacharjee, 2014; Emamian et al., 2014; Fenske et al., 2013; Mostafa, 2011; Rannan-Eliya et al., 2013; Saha et al., 2009; Srinivasan et al., 2013).

In summary, the literature on factors influencing under-five children's nutrition has shown that indeed it is the result of multiple factors that influence the nutritional outcomes of children and that such factors could be operating at the individual child, household or community-levels and the identification of these factors is very crucial for the fuller understanding and better targeting of the problem. Identification of factors responsible for childhood malnutrition could lead to policy formulation and implementation for governmental and non-governmental organizations. Also, addressing the problem of childhood malnutrition is an ongoing process and it is therefore pre-requisite in this process to conduct frequent nutritional studies to investigate the prevalence and its determinants.

1.6 The need for this study

Childhood malnutrition is a chronic problem and one of the major public health challenges in Ghana. Several attempts from the government of Ghana and non-governmental organizations (e.g. UNICEF, Food and Agriculture Organization, the World Food Programme, etc.) to address the problem have fallen below expectation. Some of the attempts from government and non-governmental organizations aimed at addressing child malnutrition include: 1) Early initiation and intensification of exclusive breastfeeding and complementary feeding of children aged below 5 years; 2) management of child malnutrition (e.g. vitamin A supplementation); 3) management of the sick newborn; 4) growth monitoring and promotion using charts based on WHO Growth Standard; 5) prenatal nutrition including iron and foliate supplementation and 6) detection and treatment of problems complicating pregnancy. The Government of Ghana has supported the formulation and implementation of important legislation and policies with specific targets for nutrition: 1) Growth and Poverty Reduction Strategy I and II; 2) Ghana Shared Growth and Development Agenda I and II; 3) Breastfeeding Promotion and Regulation; 4) Infant and Young Child Feeding Strategy and 5) Vitamin A Policy (Government of Ghana, 2013).

According to the 2013 Ghana National Nutrition Policy for 2014-2017, the government of Ghana recognised the fact that “malnutrition is caused by a wide array of factors, which must be identified, prioritized, and addressed” (Government of Ghana, 2013). The Government of Ghana (GoG) also recognised the fact that desired outcomes in nutrition in Ghana have not been achieved for several reasons: 1) nutrition and

nutrition-related interventions implemented by various sectors have not been adequately prioritized, coordinated, and integrated; 2) nutrition has not been prioritized as a key development issue and thus has not received adequate political and financial investments; 3) the sheer scope of the problem is enormous-the entire population, especially women and children suffer from all the major micronutrient deficiencies (Government of Ghana, 2013). The GoG also recognised that the slow progress in addressing poor child feeding practices, infections, and food insecurity deterred progress in reducing malnutrition in the country (Government of Ghana, 2013).
This study therefore seeks to fill some of the knowledge gaps expressed by GoG. Below are specific knowledge gaps this study attempts to fill with the hope of helping the GoG and NGOs to facilitate actions aimed at minimizing child malnutrition in the country:

(1) Limited studies are available indicating determinants of malnutrition among children. As a result, less has been published about the social, economic and environmental factors facing the residents of Ghana that may explain the country's relatively high rates of childhood malnutrition and, consequently, childhood and maternal mortality. This poses a key risk-factor knowledge gap for policymakers as this is necessary for planning effective policies, interventions and programmes aimed at improving childhood nutrition. This thesis seeks to fill this gap by examining major risk factors for childhood nutrition in Ghana and to further examine whether or not the nutritional outcome of children in Ghana differ substantially across households.
(2) Most of the limited studies (Antwi, 2008; Ghana Statistical Service et al., 2009) conducted in Ghana to examine child nutrition situations in the country did not use appropriate statistical methods in their analysis, especially those attempting to identify risk factors for child malnutrition. This could substantially impact

negatively on the interpretation of the findings and, consequently, inefficient policies, interventions and programmes and this thesis seeks to fill this gap.

(3) Childhood malnutrition could vary geographically across Ghana but an appropriate study to examine the geographical variations in the risk of under-five children's malnutrition continuously across the whole of Ghana to inform appropriate and efficient nutrition policies and interventions in the higher risk communities to be identified is presently unavailable in the country. This is key information gap facing policymakers responsible for the nutrition and health of children in Ghana because the public health resources available in the country are very limited and should be spent on the higher risk communities rather than universal intervention which is presently not feasible in the country. The aim of examining the geographical differences in the nutritional outcomes of children across Ghana in this thesis is to help policymakers in the targeting of scarce resources to the communities that needed it most and in the optimisation of public health policy intervention strategies in the country.

(4) Nutritional outcome of children could vary over place and time but none of the studies in Ghana investigates spatial and temporal variations in childhood chronic malnutrition. This is crucial because to better inform carefully targeted interventions to reduce childhood malnutrition prevalence to some appreciable level in Ghana, public health planners and policymakers need access to timely and relevant malnutrition prevalence data, trend analyses over time and place, and forecast estimates but these are largely unavailable presently. Also, since malnutrition prevalence persist in Ghana, continued examination of the trend in the risk of malnutrition over time and space and their confounders is warranted if an appreciable level of progress is to be made in reducing the prevalence

closer to those expected in a healthy, well-fed under-five children population. This study seeks to fill this gap by investigating spatial and temporal trends in the risk of malnutrition and to identify communities at highest risk overtime in Ghana.

(5) The present childhood malnutrition interventions in Ghana are reactive: as child malnutrition prevalence persists in Ghana, mass interventions are conducted but this does not cover everyone affected due to limited public health resources in the country. Thus, this approach will not result in any real solution to the problem because the mass intervention might not reach those affected most or the communities at higher risk of malnutrition. This study seeks to feel this gap by identifying communities where the risk of malnutrition is highest to enable efficient and targeted nutrition policy and interventions in such communities to minimise the risk and to provide forecast estimates to inform pre-emptive malnutrition interventions, prioritisation, and effective and sustainable public health nutrition policies in Ghana.

(6) Also, addressing the problem of childhood malnutrition is an ongoing process and it is therefore pre-requisite in this process to conduct frequent nutritional studies to investigate the prevalence and its determinants.

The findings from this thesis are intended to help policymakers responsible for the health and nutrition of children to design efficient public health policies and targeted nutrition interventions and programmes amidst scarce public health resources available in Ghana to better understand, target and to minimise childhood malnutrition prevalence closer to the level expected in a healthy, well-fed population of children under-fives.

1.7 Study objectives, research questions, contributions and organization

The work in this thesis is motivated by the persistent childhood malnutrition prevalence, one of the major real-life global public health problems especially in developing countries like Ghana. We aim to apply appropriate and advanced statistical methods to this problem to better inform efficient public health policies and malnutrition interventions. The study objectives and research questions in this thesis are presented in Table 1.3.

Table 1.3 Study objectives and research questions

Study objectives	Research questions
To investigate the major risk factors for malnutrition to be assessed on weight-for-age, height-for-age and weight-for-height Z-scores among children in Ghana	What are the major risk factors for under-five child malnutrition in Ghana?
To investigate household level variations in nutritional outcomes of children to be measured on weight-for-age, height-for-age and weight-for-height Z-scores	Does childhood nutritional outcome vary across households in Ghana?
To explore, model and illustrate spatial variation in the risk of childhood malnutrition to be assessed on weight-for-age, height-for-age and weight-for-height Z-scores among children in Ghana	Does the risk of malnutrition in children vary spatially in Ghana?
To explore, model, forecast and illustrate spatio-temporal variation in the risk of childhood malnutrition to be assessed on height-for-age Z-scores among children in Ghana	Does the risk of childhood chronic malnutrition among children in Ghana vary over place and time?
To model weight-for-age (WAZ) and height-for-age (HAZ) Z-scores jointly while adjusting for risk factors	What is the extent to which the residual correlation effects for WAZ and HAZ scores depend on children and households? Does modelling WAZ and HAZ scores jointly improves the accuracy and reliability of parameter estimates? How does the important risk factor effects compare across WAZ and HAZ scores?

The statistical methods we considered in this thesis cover methods for analysing multilevel, multivariate response, spatial and spatiotemporal data. Presented below are the contributions provided by this thesis:

- Chapter 4: investigates the major risk factors associated with childhood malnutrition and further examines variations in childhood nutritional outcomes across households. This study is crucial because malnutrition prevalence persists in Ghana and continued examination of the malnutrition prevalence and its associated risk factors using appropriate and advanced statistical methods is warranted if an appreciable level of progress is to be made in reducing the prevalence closer to those expected in a healthy, well-fed population of children under-fives. Furthermore, investigating household-level differences in childhood nutrition allows identification of any unanswered questions about the reasons for variations in malnutrition associated with the household in which a child resides which could inform further studies to identify as-yet unidentified factors which might account for the unexplained household-level variations in childhood nutritional outcomes.
- Chapter 5: our primary aim in this study is to explore, model and illustrate spatial variations in the risk of under-fives malnutrition and to identify communities where the risk of malnutrition are highest. This study is important because it can help in the development of an efficient public health intervention strategy to better understand, target and address the problem of childhood malnutrition in the high risk communities amidst the limited public health resources available in most developing countries like Ghana. This study is also the first to apply appropriate and advanced spatial modelling methods to investigate spatial variations in childhood malnutrition over Ghana and among the few studies in developing countries that investigate spatial variations in childhood nutritional outcomes.

- Chapter 6: we explore, model, forecast and illustrate spatio-temporal variations in the risk of under-fives stunting (chronic malnutrition) and identify communities where the risk of stunting are highest based on our forecast. To better inform carefully targeted interventions to reduce childhood malnutrition, public health planners and policymakers need access to timely, relevant malnutrition prevalence data, trend analyses and forecast estimates but these are largely unavailable presently in Ghana. The importance of this study is based on the fact that it will enable pre-emptive malnutrition interventions, prioritisation, effective and sustainable public health nutrition policies in Ghana, as the saying goes 'prevention is better than cure'. Results from this study can be used to target public health nutrition interventions so as to improve overall health and nutrition of Ghanaian children. In addition, this study is the first in developing countries to apply novel statistical models to forecast spatio-temporal variation in the risk of under-fives stunting prevalence.
- Chapter 7: we examine the degree to which the residual correlations between nutritional outcomes measured on weight-for-age (a measure of underweight) and height-for-age (a measure of stunting) Z-scores depend on children and their households, examine the major risk factors for under-fives nutrition as well as a specific effect of a risk factor across these outcomes. We also examine the variance explained by households in WAZ and HAZ scores among children. This study is crucial because in the presence of at least a moderate empirical correlation between the outcomes, modelling these outcomes jointly will allow inference about the degree to which the residual correlations between the outcomes depend on children and their households. For example, a strong residual correlation from the joint models, either on children or households or

both, will suggest that as-yet unidentified factors (eg. household food availability, maternal health seeking behaviour, feeding practices) at child or household-level or both that influence both WAZ and HAZ scores are the same. This could help inform policy decisions since any intervention that could investigate and address such as-yet unidentified factors influencing WAZ score can equally do for HAZ score. Thus, there may be no need to develop separate interventions to address the factors at the child or household-level or both influencing both WAZ and HAZ scores. This study will also help improve accuracy and reliability in estimates of the model parameters as opposed to modelling the outcomes separately.

Each of the papers presented in Chapters 4, 5, 6 and 7 include their own abstract, introduction, method, result, discussion, conclusion and reference sections and hence can be read separately. The thesis also provides information on the data and the general overview of statistical methods in Chapters 2 and 3 respectively, as well as general discussion and future work in Chapter 8.

This thesis focuses on modelling childhood malnutrition among children under-five population using multilevel, multivariate response, spatial and spatio-temporal data. As a result, we reviewed the literature on analysing multilevel, multivariate response, spatial and spatio-temporal data in chapter 3 of this thesis with an emphasis on statistical modelling. Readers who might be interested in other aspects can refer to the books we cited.

1.8 Role in the published/submitted/non submitted works

Presented below is the role of Justice Moses Kwaku Aheto in the published/submitted/non submitted works presented in Chapters 4, 5, 6 and 7 as follows:

- Chapter 4: Conceived the idea of the study, data acquisition and preparation, statistical analysis, preparation of manuscript, and final version of the paper.
- Chapter 5: Conceived the idea of the study, data acquisition and preparation, statistical analysis, and preparation of manuscript
- Chapter 6: Conceived the idea of the study, data acquisition and preparation, statistical analysis, and preparation of manuscript.
- Chapter 7: Conceived the idea of the study, data acquisition and preparation, statistical analysis and preparation of brief report.
- Furthermore, presented below are the roles of co-authors in the published/submitted papers presented in Chapters 4, 5 and 6:
- Benjamin M. Taylor contributed to the analyses, interpretation and critically reviewed the manuscripts. Thomas J. Keegan contributed to the interpretation and critically reviewed the manuscripts and Peter J. Diggle contributed to the interpretation and critically reviewed the manuscripts. All the authors read and approved the final draft of the manuscripts and agreed to be accountable for all aspects of the work in these papers.

Reference

Adekanmbi, V. T., G. A. Kayode, and O. A. Uthman (2013). Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis: Matern Child Nutr, v. 9, p. 244-59.

Alom, J., M. A. Quddus, and M. A. Islam(2012). Nutritional status of under-five children in Bangladesh: a multilevel analysis: J Biosoc Sci, v. 44, p. 525-35.

Antwi, S. (2008). Malnutrition: missed opportunities for diagnosis: *Ghana Medical Journal*, v. 4**2**, p. 101-104.

Babatunde, R. O., F. I. Olagunju, S. B. Fakayode, and F. E. Sola-Ojo (2011). Prevalence and Determinants of Malnutrition among Under-five Children of Farming Households in Kwara State, Nigeria: Journal of Agricultural Science, v. 3.

Black R. E., C.G. Victoria, S. P. Walker and the Maternal and Child Nutrition Study Group (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. Lancet 2013: published online June 6, available at http://dx.doi.org/10.1016/S0140-6736(13)60937-X

Bhutta ZA, Das JK, Rizvi A, et al, The Lancet Nutrition Interventions Review Group, and the Maternal and Child Nutrition Study Group. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? *Lancet* 2013; published online June 6. http://dx. doi.org/10.1016/S0140-6736(13)60996-4.

Bomela, N. J. (2009). Social, economic, health and environmental determinants of child nutritional status in three Central Asian Republics: Public Health Nutr, v. 12, p. 1871-7.

Dang, S., H. Yan, S. Yamamoto, X. Wang, and L. Zeng (2004). Poor nutritional status of younger Tibetan children living at high altitudes: Eur J Clin Nutr, v. 58, p. 938-46.

Debnath, A., and N. Bhattacharjee (2014). Factors associated with malnutrition among tribal children in India: a non-parametric approach: J Trop Pediatr, v. 60, p. 211-5.

Emamian, M. H., M. Fateh, N. Gorgani, and A. Fotouhi (2014). Mother's education is the most important factor in socio-economic inequality of child stunting in Iran: Public Health Nutr, v. 17, p. 2010-5.

Engebretsen, I. M., T. Tylleskär, H. Wamani, C. Karamagi, and J. K. Tumwine (2008). Determinants of infant growth in Eastern Uganda: a community-based cross-sectional study: BMC Public Health, v. 8, p. 418.

Fenske, N., J. Burns, T. Hothorn, and E. A. Rehfuess (2013). Understanding child stunting in India: a comprehensive analysis of socio-economic, nutritional and environmental determinants using additive quantile regression: PLoS One, v. 8, p. e78692.

Food and Agriculture Organization (FAO) (2009). Nutrition Country Profiles-Ghana. [last accessed May 2015] (http://www.fao.org/ag/AGN/nutrition/gha_en.stm), Rome, FAO, United Nations.

Ghana Statistical Service (2012). *Population and housing census: summary report of final results 2010*, Accra, Ghana., GSS.

Ghana Statistical Service, Ghana Health Service, and ICF Macro (2009). Ghana Demographic and Health Survey 2008, Accra, Ghana., GSS, GHS, ICF Macro.

Government of Ghana (GoG) (2013). National Nutrition Policy 2014-2017, Accra, Ghana. [last accessed September 2015] (https://extranet.who.int/nutrition/gina/sites/default/files/GHA%202013%20National%20Nutrition%20Policy.pdf), Accra, GoG.

Kabubo-Mariara, J., G. K. Ndenge, and D. K. Mwabu (2009). Determinants of Children's Nutritional Status in Kenya: Evidence from Demographic and Health Surveys.: Journal of African Economies, v. 18, p. 363–387.

Kandala, N. B., T. P. Madungu, J. B. Emina, K. P. Nzita, and F. P. Cappuccio (2011). Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter?: BMC Public Health, v. 11, p. 261.

Kinyoki, D. K., J. A. Berkley, G. M. Moloney, N. B. Kandala, and A. M. Noor (2015). Predictors of the risk of malnutrition among children under the age of 5 years in Somalia: Public Health Nutr, p. 1-9.

Madise, N., and M. Mpoma (1997). Child malnutrition and feeding practices in Malawi, v. 18, p. 190-201.

Mbago, M. C., and P. P. Namfua (1992). Some determinants of nutritional status of one-to four-year-old children in low income urban areas in Tanzania: J Trop Pediatr, v. 38, p. 299-306.

Ministry of Food and Agriculture (MoFA) (2013). Agriculture in Ghana: Facts and Figures (2012), Accra, Ghana, MoFA.

Mostafa, K. S. (2011). Socio-economic determinants of severe and moderate stunting among under-five children of rural Bangladesh: Malays J Nutr, v. 17, p. 105-18.

Pelletier DL, Frongillo EA, and Habicht JP (1993). Epidemiological evidence for a potentiating effect of malnutrition on child mortality: *American journal of Public Health*, v. **83**, p. 1133-1139.

Rannan-Eliya, R. P., S. M. Hossain, C. Anuranga, R. Wickramasinghe, R. Jayatissa, and A. T. Abeykoon(2013). Trends and determinants of childhood stunting and underweight in Sri Lanka: Ceylon Med J, v. 58, p. 10-8.

Ruel MT, Alderman H, and the Maternal and Child Nutrition Study Group. Nutrition-sensitive interventions and programmes: how can they help to accelerate progress in improving maternal and child nutrition. *Lancet* 2013; published online June 6. http://dx.doi.org/10.1016/S0140- 6736(13)60843-0.

Saha, K. K., E. A. Frongillo, D. S. Alam, S. E. Arifeen, L. A. Persson, and K. M. Rasmussen (2009). Household food security is associated with growth of infants and young children in rural Bangladesh: Public Health Nutr, v. 12, p. 1556-62.

Srinivasan, C. S., G. Zanello, and B. Shankar (2013). Rural-urban disparities in child nutrition in Bangladesh and Nepal: BMC Public Health, v. 13, p. 581.

The World Bank (2006). Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action, Washington DC, The International Bank for Reconstruction and Development/The World Bank.

UNICEF (1990). Strategy for improved nutrition of children and women in developing countries: a UNICEF policy review, New York, UNICEF, p. 22.

UNICEF, WHO, The World Bank (2015). Joint child malnutrition estimates - Levels and trends: Key findings of the joint malnutrition estimates (2015 edition). Accessed on 10/12/2015 at http://www.who.int/nutgrowthdb/estimates2014/en/

Vella, V., A. Tomkins, A. Borghesi, G. B. Migliori, B. C. Adriko, and E. Crevatin (1992). Determinants of child nutrition and mortality in north-west Uganda: Bull World Health Organ, v. 70, p. 637-43.

Wamani, H., A. N. Astrøm, S. Peterson, J. K. Tumwine, and T. Tylleskär (2006). Predictors of poor anthropometric status among children under 2 years of age in rural Uganda: Public Health Nutr, v. 9, p. 320-6.

Chapter 2 The Ghana Demographic and Health Survey Data

2.1 Introduction

The aim of this chapter is to describe and discuss the data used for the analyses in this thesis, the Ghana Demographic and Health Survey (GDHS).

The GDHS is a series of national population and health surveys conducted in Ghana every five (5) years to provide current and reliable information on maternal and child health, marriage, breastfeeding practices, sexual activity, fertility preferences, fertility levels, awareness and use of family planning methods, nutritional status of women and young children, childhood mortality, domestic violence, and awareness and attitudes regarding AIDS and other sexually transmitted infections. The GDHS is part of the wider Demographic Health Program (DHS) which aims to collect data to be used to monitor and evaluate population health and nutrition programmes, and does so in over 90 countries (MEASURE DHS, 2016, MEASURE DHS 2015).

2.2 Survey designs

The GDHS shares a common set of design principles with the DHS surveys (MEASURE DHS, 2015; MEASURE DHS, 2014). Standard DHS surveys take place usually every 5 years and have sample sizes between 5,000 and 30,000 households. The sample is representative at the national, residence (urban-rural) and regional levels. DHS surveys are normally conducted over a period of 18-20 months.

In common with other DHS surveys, the GDHS uses a two-stage cluster sampling design for each survey. We used data from four surveys: those conducted in 1993, 1998, 2003 and 2008. The 1984 Ghana Population Census (Ghana Statistical Service, 1987) provided the master sampling frame from which the clusters or the sampling points were selected for the 1993 and 1998 GDHSs whilst the 2000 Ghana Population and Housing Census (GPHC) (Ghana Statistical Service, 2002) provided the master sampling frame for the selection of the clusters for the 2003 and 2008 GDHSs. The clusters were selected with probability proportional to the number of households listed in each cluster. A total of 400, 400, 412 and 411 clusters were used in the 1993, 1998, 2003 and 2008 GDHSs, respectively. The selection of the clusters was followed by a complete listing of all households in the selected clusters, which provided the sampling frame for the selection of households in the second stage. Different samples and sample sizes were used in each of the four surveys, so it was a serial cross-sectional rather than a longitudinal design.

The number of households sampled for interviews in 1993, 1998, 2003 and 2008 are 6,161, 6,375, 6,628 and 11,913, respectively. Each survey used the same set of three questionnaires: (1) a household questionnaire (HQ), (2) a women's questionnaire (WQ) and (3) a men's questionnaire (MQ). Women were aged 15-49 years (i.e. women of reproductive age) and men were aged 15-59 years. The HQ was used to list all the usual members and the visitors in the selected households. Some basic characteristics such as age, education, sex, and relationship with head of household was collected on each person listed. The main purpose of the HQ was to identify women and men who were eligible for individual interview. The WQ was used to obtain information from women aged 15-49 years in half of the selected households and the selected women were also asked questions about their children aged below 5 years. Some of the information collected using WQ include education, knowledge and use of family planning methods, breastfeeding and infant and young child feeding practices, anthropometric measurements on women and their children, sexual activity, child mortality, awareness and behaviour about AIDS and other sexually transmitted infections. The MQ on the other hand was used administered to men aged 15-59 in half of the households selected. The MQ collected nearly same information collected by WQ except that it did not ask questions about reproductive history or questions on maternal and child health or nutrition.

Presented in Table 2.1 are the periods of interview for the four data surveys. This is important because the period of the year in which people are asked about nutrition can affect their answers (introduction of bias). This is because food availability varies seasonally, with food shortages during what is known as the hunger season". In Ghana this last from around February to June in the northern part of the country and from around April to June in the southern part of the country.

Table 2.1 Time of interview period for data sets used in this thesis

Survey	Interview period	Season (hunger/no hunger)
1993 GDHS	September 1993 to February 1994	not a hunger season
1998 GDHS	November 1998 to February 1999	not a hunger season
2003 GDHS	July 2003 to October 2003	not a hunger season
2008 GDHS	September 2008 to November 2008	not a hunger season

2.3 Data sets used in this thesis

Presented in Figure 2.1 is the structure of the data sets with clusters located at the top in which households are nested. Below the clusters is the number of households sampled for interview and below the households is the number of children under 5 years of age selected from all households from which female individual interviews were conducted. The household level response rates for the 1993, 1998, 2003 and 2008 GDHS is 98.4%, 99.1%, 98.7% and 98.9% respectively (Ghana Statistical Service et al., 2009; Ghana Statistical Service et al., 2004; Ghana Statistical Service (GSS) and Macro International Inc (MI), 1994, 1999). The 2003 GDHS has the largest number of children aged below 5 years whilst the 1993 GDHS has the smallest number of children below 5 years of age.

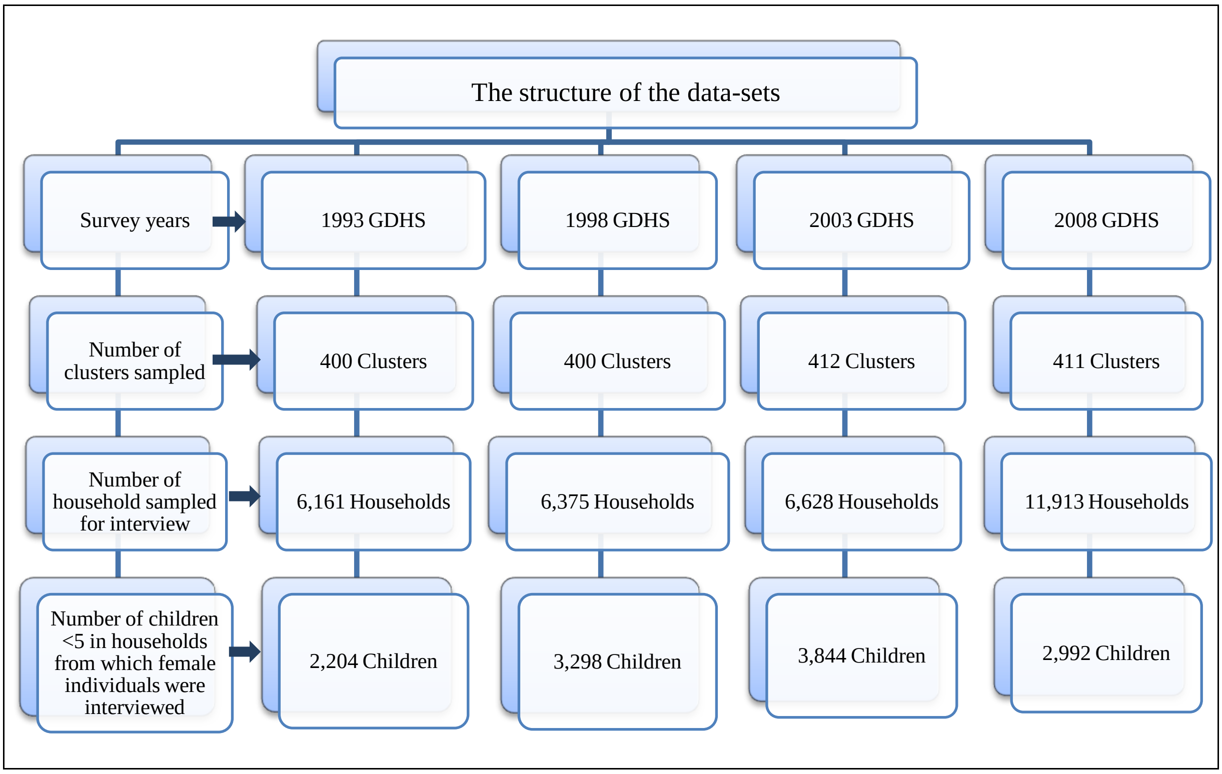


Figure 2.1 The hierarchical structure of the data-sets used in this thesis

2.3.1 Data quality and limitations

Data obtained from Demographic and Health Surveys (DHS) such as GDHSs are recognized for their accuracy and source of nationally representative information detailing the demographic and health status of developing countries like Ghana with much focus on young children and their mothers.

The four GDHSs used in this thesis contained good quality data on a number of child and maternal/household characteristics as well as data on geographic information. Data quality is of utmost importance to the DHS Program which is considered the gold standard for nationally representative data collection in developing countries (Kiersten, 2009). The following were some of the measures put in place to ensure data quality in the GDHSs: 1) field officers (interviewers, editors and supervisors) and data entry personnel are trained extensively on the survey instruments. The training which is supported with interviewer and supervisor manuals involves classroom discussions and practice focusing on the three questionnaires (Household Questionnaire (HQ), Women's Questionnaire (WQ) and Men's Questionnaire (MQ)) based on model questionnaires developed by the MEASURE DHS programme; 2) Trainees selected as field editors and supervisors were given additional days of training on how to edit questionnaires and supervise fieldwork; 3) Training of interviewers was conducted mostly in English by senior staff from the Ghana Statistical Service (GSS) with technical support from ICF Macro; 4) A pre-test is conducted to test the survey instruments and to identify possible challenges ahead of the main survey; 5) Only female interviewers interviewed respondents for the WQ and only male interviewers interviewed respondents for the MQ; 6) To reduce language barrier, the questionnaires were translated from English into three (3) major local languages, namely Ewe, Twi and Ga; 7) Interviewers were selected on the basis of in-class participation, field practice, fluency in the Ghanaian languages, and an assessment test. The most experienced trainees, those who had participated in the pre-test and those who did extremely well during the training were selected to be supervisors and editors; 8) Senior staffs from GSS coordinated and supervised the field work whereas officials from ICF Macro participated in field supervision of interviews. The main reason for their visits is to check the quality of data being collected, ascertain whether the right procedures were being followed and assist in resolving any challenges that a team might be experiencing; 9) Completed questionnaires were returned periodically from the field to the GSS office in Accra where they were entered and edited by data processing personnel specially trained for this task. The concurrent entry of the data gave the GSS the opportunity to advise field teams of problems detected during data entry which has an added advantage for data quality; 10) All data were entered twice for 100% verification (Ghana Statistical Service et al., 2009; Ghana

Statistical Service et al., 2004; Ghana Statistical Service (GSS) and Macro International Inc (MI), 1994, 1999).

The GDHS collect information on similar variables in each survey iteration and these variables are named in the same way across all survey time periods. The GDHS also used the same survey methodology and were conducted by the same organization.

In this thesis, we used data on children aged below 5 years with eligible anthropometric measurements (eg. weight and height) from the women's questionnaire in each of the four GDHSs. We computed the weight-for-age, height-for-age and weight-for-height Z-scores using the 2006 World Health Organization growth standards and used these as measures of child's nutritional outcomes in this thesis. The Z-scores used in the analysis are all within the plausible ranges and available for all children in this thesis.

All information on children in the GDHS data was reported by the mother apart from those that required measurement and test. For instance, information on child's diarrhoea episodes is based on mother's understanding of the signs and symptoms and this understanding could vary from one mother to another or can differ across settings. This could introduce reporting bias. Also, it is difficult to directly measure household wealth status in Ghana. Because of this we used an asset-based index, which is generally considered a good proxy for household wealth status in developing countries. Another limitation of the data is incomplete information on child's actual birthweight and as a result, the association between child's nutritional outcomes and this variable could not be explored directly.

Furthermore, though complementary feeding is one of the key covariates to be considered in our analysis, we were unable to explore the association between child's nutritional outcomes and this variable in this study due to inadequate data on complementary feeding. An alternative approach will be to undertake subsample analysis but we did not pursue this because only 19% of children in the data-set had data on complementary feeding and this proportion could further reduce when other risk factors are accounted for should subsample analysis be considered. We hope future GDHSs will collect enough data on complementary feeding to allow the association between child's nutritional outcomes and complementary feeding to be explored in further studies in this area. Other limitation includes the fact that some children in the GDHSs have unacceptable or biologically implausible Z-scores and could not be included in the analysis.

2.3.2 Variables used in the analysis

Presented in Table 2.2 is the description of dependent and candidate explanatory variables used in this thesis.

Table 2.2 Description of variables used in this thesis

Child characteristics
Dependent variables used in this thesis: Weight-for-age Z-score: a scale or measure from which underweight is derived and is continuous. Height-for-age Z-score: a scale or measure from which stunting is derived and is continuous. Weight-for-height Z-score: a scale or measure from which wasting is derived and is continuous.
Sex: 1= Male and 0= Female
Age: Age in years and is continuous
Breastfeeding duration: breastfeeding duration in months and continuous
Type of birth: 1= multiple birth and 0= single birth
Diarrhoea episodes: 1= child had diarrhoea and 0= child did not have diarrhoea
Fever episodes: 1= child had fever and 0= child did not have fever
Size at birth: 1= small and 0= average/large
Place of birth delivery: 1= not delivered at health care facility and 0= delivered at health care facility
Maternal/household characteristics
Mother's education: mother's education in years and is continuous
Mother's body mass index (BMI): mother's BMI and is continuous
Mother's age: mother's age in years and is continuous
Mother's health insurance status: 1= mother did not have health insurance and 0= mother had health insurance
Household wealth status: 1= poor and 0= average/rich
Household source of drinking water: 1= non-piped source and 0= piped source
Type of toilet facility in household: 1= no facility and 0= flush/pit facility
Number of dead siblings in household: 1= one or more and 0= none
Household size: number of persons in a household and is continuous
Community characteristics
Residence: 1= rural and 0= urban
Vegetation index (NDVI): a measure of a vegetation cover and is continuous
Elevation (km): a measure of altitude (height above sea level) and is continuous

2.3.3 Process of selecting explanatory variables used in this thesis

In all regression analyses we employed a standard backward elimination method for selecting potential explanatory variables. Under this technique, we start by fitting the model with all the potential explanatory variables of interest and testing deletion of the candidate variables one by one using a chosen model comparison criterion. We eliminated variables that were not significant at the level of 10% (0.10): each variable which is not significant at this critical level is deleted from the model, starting with the least. The process is repeated by successively refitting reduced models, applying the same rule until all the remaining variables are statistically significant at the chosen critical level (i.e. until no further improvement is possible). We used the variables retained at this stage in our final analysis (Draper, 2014). (REF)

2.3.4 Comparing the GDHS to other DHS countries

It is useful to compare the findings of the Ghana Demographic Health survey with results of DHS programs in other countries. The programs use the same study methodology, collect information on the same variables which are named in the same way across these countries. The survey manuals and the technical assistance provided ensure that the survey procedures followed in each country are similar, permitting cross-country data comparison across countries participating in the MEASURE DHS program.

Usually, data sets from the DHS surveys are made available online through a process of electronic registration and descriptive survey results are published in final reports. The data sets are also available at the offices of those responsible for the conduct of the surveys in the participating countries (MEASURE DHS, 2014).

Table 2.3 shows the sample sizes and response rates from selected DHS surveys. The number of households interviewed varied by country, but in most countries the response rate was over 90% for the men and women asked for interview, though men's response rates were always slightly lower than for women.

Table 2.3 Summary results of some selected DHS 2008-2011

Country	Survey year	Number of households interviewed	Household response rate	Number of women interviewed	Women response rate	Number of men interviewed	Men's response rate
Ghana	**2008**	**11,778**	**98.9**	**4,916**	**96.5**	**4,568**	**95.8**
Nigeria	2008	34,070	98.3	33,385	96.5	15,486	92.6
Egypt	2008	18,968	99.1	16,527	99.7	NA	NA
Guyana	2009	5,632	93.2	4,996	90.1	3,522	77.4
Armenia	2010	6,700	95.1	5,922	97.7	1,584	96.5
Tanzania	2010	9,623	98.8	10,139	96.4	2,527	91.2
Rwanda	2010	12,540	99.8	13,671	99.1	6,329	98.7
Bangladesh	2011	17,141	97.9	17,842	97.9	3,997	92.0
Ethiopia	2011	16,702	98.1	16,515	95.0	14,110	88.7
Uganda	2011	9,033	95.3	8,674	93.8	2,295	89.2

NA = Not Available

Reference

Draper, N. R. (2014). Backward Elimination Selection Procedure. Wiley StatsRef: Statistics Reference Online.

Ghana Statistical Service (1987). *1984 Population Census of Ghana: Demographic and Economic Characteristics (II volumes: Total Country, Western, Central, Greater Accra, Eastern, Volta, Ashanti, Brong-Ahafo, Northern, Upper West and Upper East regions)*. Accra, Ghana.

Ghana Statistical Service (2002). *2000 Population and Housing Census, summary report of final results*. Accra, Ghana.

Ghana Statistical Service, Ghana Health Service, and ICF Macro (2009). Ghana Demographic and Health Survey 2008, Accra, Ghana., GSS, GHS, ICF Macro.

Ghana Statistical Service, Noguchi Memorial Institute for Medical Research, and ORC Macro (2004). Ghana Demographic and Health Survey 2003, Calverton, Maryland, GSS, NMIMR, and ORC Macro.

Ghana Statistical Service (GSS), and Macro International Inc (MI) (1994). Ghana Demographic and Health Survey 1993, Calverton, Maryland, GSS and MI.

Ghana Statistical Service (GSS), and Macro International Inc (MI) (1999). Ghana Demographic and Health Survey 1998, Calverton, Maryland, GSS and MI.

Kiersten J, Grant M, Khan S, Moore Z, Armstrong A, and Sa Z. (2009). Fieldwork-Related Factors and Data Quality in the Demographic and Health Surveys Program. DHS Analytical Studies No. 19. Calverton, Maryland, USA: ICF Macro.

MEASURE DHS. (2014). The DHS Program: Datasets [Assessed on 15/06/2014 at http://www.measuredhs.com/data/available-datasets.cfm

MEASURE DHS. (2015). *DHS Survey Types: Standard DHS Surveys.* The DHS Program. [Accessed on 22/11/2015, available at https://dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm].

MEASURE DHS. (2016). The DHS Program [Accessed on 18/09/2016, available at http://dhsprogram.com/].

Chapter 3 General Overview of Statistical Methods

This chapter presents the general overview of statistical methods employed to answer the research questions presented in section 1.7 in Table 1.3.

3.1 Multilevel and multivariate response multilevel modelling

This section briefly presented the description of the general multilevel and multivariate response multilevel modelling methods.

3.1.1 Multilevel modelling

Multilevel modelling technique (Goldstein, 2003; Snijders and Bosker, 2012) is a methodology for analysing data which are clustered, hierarchical or nested in nature. The complex nature of these data introduces different sources of variability and this section of the thesis focuses on nested source of such variability. For instance, patients nested within doctors. Thus, there is a different source of variability between patients and between doctors and not recognizing such distinctively could lead to incorrect conclusions. In analysing such data, it is crucial to take into account the various levels of nesting because they might be associated with variability which has a distinct interpretation.

Multilevel analysis are been used in many disciplines such as behavioural, social, biomedical and life sciences. The paper by Robinson discussed the need to recognise individual and group level effects distinctively and provided detailed discussion on the mix-up between aggregate and individual level effects, termed as ecological fallacy (Robinson, 1950) and Davis' et al. also discussed the distinction between within-group and between-group regression analysis (Davis et al., 1961). For detailed history about multilevel analysis, see the bibliographical sections of Longford (Longford, 1993).

Let *i* denotes level-one unit (e.g., individual) and *j* denotes level-two unit (e.g., the group an individual belongs to). Also, let Y_{ij} indicates an outcome variable for individual *i* in group *j* and x_{ij} be explanatory variable for child *i* in group *j* which may be directly related to the individual or the group the individual belongs to. A two-level random intercept variance components model in a multilevel framework can be presented as:

$$Y_{ij} = \beta_0 + U_{0j} + \beta_1 x_{ij} + \varepsilon_{ij} \quad (3.1)$$

with U_{0j} ~ N(0, σ^2_u), and ε_{ij} ~ N(0, σ^2_ε). The regression coefficient β_1 is common to all the groups, β_0 is the average intercept, U_{0j} is the group-dependent deviation and ε_{ij} is an individual-level residual.

The variance components σ^2_u and σ^2_ε are used to obtain the intra-class correlation coefficients (ICC) which coincides with variance partition coefficients (VPC) for the model (3.1) given as $\{\sigma_u^2/(\sigma_u^2 + \sigma_\varepsilon^2)\}$ x 100, a measure of the amount of variation (%) explained by the group. This model (3.1) can be extended to more than 2 levels when more levels of grouping are made available.

We can also extend model (3.1) by allowing the explanatory variable x_{ij} to have a random coefficients at a level higher than the individual level, where the variation of a group-level factor might differ significantly across the groups. This model is presented as:

$$Y_{ij} = \beta_0 + U_{0j} + \beta_1 x_{ij} + U_{1j} x_{ij} + \varepsilon_{ij} \quad (3.2)$$

The group-dependent coefficients (U_{0j} ,U_{1j}) are assumed to be independent across j with a bivariate normal distribution. They have (0, 0) expected values, and covariance matrix given by:

var(U_{0j}) = σ^2_{u0}; var(U_{1j}) = σ^2_{u1}; and cov(U_{0j}, U_{1j}) = σ_{u01}. For further discussion on multilevel modelling techniques, see (Goldstein, 1987; Goldstein, 2003; Snijders and Bosker, 2012).

3.1.2 Multivariate response multilevel modelling

Multivariate response data arises when 2 or more dependent variables (outcomes) are collected on the same individual. Note that the term ‘multivariate’ as used in this thesis refers to 2 or more outcome variables. The multivariate version of the multilevel models arises when the individual for whom 2 or more outcomes were collected belongs to a group, say examination scores on mathematics and biology (level-one units) nested on students (level-two units) within schools (level-three units).

Multivariate response multilevel models (Snijders and Bosker, 2012; Thum, 1997) could be necessary when one is interested in 2 or more outcomes measured on individuals within group and the researcher is interested in drawing conclusions about the degree to which the residual correlations depend on the individual and the group level; investigate specific effect of a covariate across 2 or more outcomes; and interested in conducting a single test of a joint effect of a covariate on 2 or more outcomes. This approach also leads to higher accuracy and reliability in estimates compared to modelling these outcomes separately, especially when these outcomes are at least moderately correlated.

A three-level multivariate response multilevel model with outcomes Y^1_{ij} and Y^2_{ij}, where the superscripts 1 and 2 denote the first and second outcome variables for an individual i within group j is of the form:

$$Y_{ij}^{1} = \beta^{(1)}X_{ij} + U_{j}^{(1)} + \varepsilon_{ij}^{(1)}$$

$$Y_{ij}^{2} = \beta^{(2)}X_{ij} + U_{j}^{(2)} + \varepsilon_{ij}^{(2)} \qquad (3.3)$$

$$\text{with, } \begin{pmatrix} U_{j}^{(1)} \\ U_{j}^{(2)} \end{pmatrix} \sim MVN\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u(1)}^{2} & \\ \sigma_{u(1,2)} & \sigma_{u(2)}^{2} \end{pmatrix}\right\}$$

$$\begin{pmatrix} \varepsilon_{ij}^{(1)} \\ \varepsilon_{ij}^{(2)} \end{pmatrix} \sim MVN\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon(1)}^{2} & \\ \sigma_{\varepsilon(1,2)} & \sigma_{\varepsilon(2)}^{2} \end{pmatrix}\right\}$$

where, X_{ij} is the covariate that can be defined at the child or household levels; $\beta^{(1)}$ and $\beta^{(2)}$ are vector of regression coefficients for Y_{ij}^{1} and Y_{ij}^{2} respectively. The quantities $\varepsilon_{ij}^{(1)}$ and $\varepsilon_{ij}^{(2)}$ are the residuals at the individual level for Y_{ij}^{1} and Y_{ij}^{2} respectively and the quantities $U_{j}^{(1)}$and $U_{j}^{(2)}$ are the residuals at the group level for Y_{ij}^{1} and Y_{ij}^{2} respectively. In addition, the quantities $\sigma^{2}{}_{u(1,2)}$ and $\sigma^{2}{}_{\varepsilon(1,2)}$ are the covariance at the group and individual levels respectively for Y_{ij}^{1} and Y_{ij}^{2}. Note that there is no level 1 variation specified because level 1 exists only to define the multivariate structure.

The corresponding population correlation coefficients (residual correlations) at the group (ρ_2) and individual (ρ_1) level are, respectively,

$$\rho_2 = \frac{\sigma_{u(1,2)}}{\sqrt{(\sigma_{u(1)}^{2} \times \sigma_{u(2)}^{2})}}, \quad and\ \rho_1 = \frac{\sigma_{\varepsilon(1,2)}}{\sqrt{(\sigma_{\varepsilon(1)}^{2} \times \sigma_{\varepsilon(2)}^{2})}}$$

It is also possible to estimate correlations between observed outcomes Y_{ij}^{1} and Y_{ij}^{2} between individuals which can be presented as:

$$\hat{\rho}(Y_{ij}^{1}, Y_{ij}^{2}) = \frac{\sigma_{u(1,2)} + \sigma_{\varepsilon(1,2)}}{\sqrt{(\sigma_{u(1)}^{2} + \sigma_{\varepsilon(1)}^{2}) \times (\sigma_{u(2)}^{2} + \sigma_{\varepsilon(2)}^{2})}}$$

Furthermore, for a hypothetical group of size n, a correlation between group means can be obtained as:

$$\hat{\rho}(\bar{Y}_j^1, \bar{Y}_j^2) = \frac{\sigma_{u(1,2)} + \sigma_{\varepsilon(1,2)}/n}{\sqrt{(\sigma_{u(1)}^2 + \sigma_{\varepsilon(1)}^2/n) \times (\sigma_{u(2)}^2 + \sigma_{\varepsilon(2)}^2/n)}}$$

Further discussion on multivariate response multilevel modelling is available elsewhere (Snijders and Bosker, 2012; Thum, 1997).

3.2 Spatial and spatio-temporal modelling

We briefly discussed the general spatial and spatiotemporal statistical methods under this section.

3.2.1 Spatial modelling

Spatial statistics is the description and analysis of spatial variations in a given measurements of interests, using spatially referenced data. In spatial epidemiology, it involves description and analysis of geographical differences in disease prevalence or risk. The interest lies in accounting for spatial autocorrelation which might exist in the data. Spatial data are common in environmental health and epidemiology, and ecology with sampling units been spatially located individuals or geographical areas (Cressie, 1993). There are different approaches to analysing spatial data:

1. Point processes; dealing with stochastic processes that generates sets of spatial points within a study region but are not concerned with any measurement made there, e.g. diarrhoea cases in a county.
2. Geostatistics; those originating from a spatially continuous process measured at finite set of data locations within a study region, e.g. rainfall.
3. Lattice data; those modelling data which are aggregated over geographical regions and are located on grids, e.g. zinc concentration in grid cells covering a

farmland. Lattice data are similar to geostatistics data except they are not continuous spatial processes but are located on grids.

This thesis mainly used geostatistical models for the spatial modelling. Generally, geostatistical models deal with phenomenon which varies over space and /or time. In geostatistical models, the measurements are modelled as a function of spatial location and other potential covariates (e.g. environmental, demographic, socioeconomic, and climatic factors). These models are often used to provide spatial predictions such as disease prevalence and pollution levels across areas or at individual locations. It also allows information on correlation structure of the modelled process or its association with covariates to be obtained through its parameter estimation framework.

The origin of geostatistics as a discipline can be traced to the mining engineer D.G Krige through his work in the mining industry in South Africa in early 1950s which involved the problem of predicting gold concentrations (Cressie, 1990). Since then, geostatistical modelling has received enormous applications in other disciplines such as soil and earth sciences, epidemiology, ecology, agriculture, microbiology, climatology and geography and its associated statistical and mathematical concept upon which it pivots as presented below.

A widely used geostatistical model is the stationary Gaussian model (Diggle et al., 2007) which is of the form

$$Y_i = \mathrm{z}(x_i^{'})\beta + \mathrm{S}(x_i) + \varepsilon_i \qquad (3.4)$$

with $\mathrm{S}(x) \sim \mathrm{N}(0, \sigma^2\rho(\mathrm{d}))$, and $\varepsilon_i \sim \mathrm{N}(0,\tau^2)$,

where, Y_i: i = 1, …, n represent the measurement values with their corresponding set of spatial locations x_i, $\mathrm{z}(x)$ is a set of geographically referenced covariates, $\mathrm{S}(x)$is the value of a correlated continuous spatial Gaussian process with zero-mean and variance

σ^2 and its associated valid correlation function ρ(d) where d is the distance between two locations given as $d = ||x - x^{'}||$, and ε_i are the independent Gaussian errors with zero-mean and variance (nugget effect) τ^2.

Commonly used valid correlation functions are Gaussian, exponential and Matérn correlation functions represented as

Gaussian: $\rho(\mathrm{d}) = \exp - (d/\emptyset)^2,$

Exponential: $\rho(\mathrm{d}) = \exp(-d/\emptyset)$

Matérn: $\rho(\mathrm{d}) = \frac{1}{2^{k-1}\Gamma(\mathrm{k})} (d/\emptyset)^{\mathrm{k}} K_k(d/\emptyset)$

where, ø > 0 is the range or scale parameter controlling the rate at which correlation decreases to zero with increasing distance, d and K_k in the Matérn correlation function is the modified Bessel function of the second kind, of order k.

One of the major interests in standard geostatistics is the technique of 'kriging' or interpolation. The goal is to predict S(x_p) at an unsampled or previously sampled location x_p. Parameter estimation based on (3.4) can be executed through maximum likelihood, given initial values provided by estimate of the variogram

$$\mathrm{V}(x, x^{'}) = \frac{1}{2} Var[S(x) - S(x^{'})]$$

and an estimate from the empirical variogram using the data Y_i: i = 1, …, n, can be defined as

$$\hat{V}_{ik} = \frac{1}{2} (Y_i - Y_k)^2, \text{ where i < k.}$$

Given the data Y_i: i = 1, …, n, the simple kriging predictor can be defined in terms of

$\hat{S}(x_p) = \Sigma_i w_i Y_i$, and w_i are chosen to minimise $Var[S(x_p) - \hat{S}(x_p)]$ subject to the constraint $E[S(x_p) - \hat{S}(x_p)] = 0$, where E denotes Expectation.

3.2.2 Spatio-temporal modelling

Spatio-temporal data originate from data sources which are associated with both space and time indexes. Examples of such data include satellite images of parts of the earth and earth quake. In epidemiology, such data sources include disease outbreaks and pollution. Most epidemiological variables normally exhibit spatial and/or spatio-temporal trends attributable to geographically varying covariates (KNOX, 1964).

The general form for spatio-temporal model can be presented as

$$Y(x,t) = S(x,t) + \varepsilon_1(x,t)$$

$$S(x,t) = m(x,t) + \varepsilon_2(x,t) \qquad (3.5)$$

where x and t, respectively indexes space and time, m is a mean process, the quantities ε_1 and ε_2 are the disturbance terms with zero-mean random processes. Given the model in (3.5), the goal is to predict the value $\hat{S}(x_p,\ t_p)$ at space-time pair $(x_p,\ t_p)$. To fit our spatio-temporal model, we made use of linear Gaussian state space model and the Kalman filtering algorithms, both of which we briefly presented below.

3.2.3 The filtering recursion

Briefly presented in this section is a discussion on filtering and the steps for computing the filtering densities under the linear Gaussian state space model. The general state space dependence structure is presented in Figure 3.1 below:

$$\begin{tikzcd}
 & Y_1 & Y_2 & & Y_{t-1} & Y_t & Y_{t+1} & & \text{(Observed)} \\
S_0 \arrow[r] & S_1 \arrow[r] \arrow[u] & S_2 \arrow[r] \arrow[u] & \dots \arrow[r] & S_{t-1} \arrow[r] \arrow[u] & S_t \arrow[r] \arrow[u] & S_{t+1} \arrow[u] & \dots \arrow[r] & \text{(Unobserved)}
\end{tikzcd}$$

Figure 3.1: State space model dependence structure

3.2.4 The Kalman Filter

This section presents the famous and an important filtering algorithm in discrete time filter theory called the Kalman filter (Kalman, 1960). Let $Y_1, Y_2 \dots, Y_{t-1,} Y_t$ be an observable time series and $S_0,\ S_1, \dots, S_{t-1}, S_t$ be an unobserved states. The Kalman filter is used when both the observation and state processes are linear with Gaussian disturbances:

$S_t = AS_{t-1} + \ BW_t$ (Unobserved/state equation)

$Y_t = \ GS_t + \ HV_t$ (Observed/measurement equation) (3.6)

The assumptions are that the state processes S_0, S_1, S_2, …, $S_{t-1,}$ S_t forms a Markov chain; the sequences W_t and V_t are independent Gaussian random variables with respective covariance W and V; W_t and V_t are uncorrelated for all t time periods; and for any given t > k, Y_t is independent of Y_k given $S_{0:k}$; and Y_t is independent of S_k for any given k ≠ t. It can be proven that the random vector (Y_1, Y_2, …, Y_t, S_1, S_2, …, S_t) has a Gaussian distribution for any t ≥ 1 through the use of standard results about the multivariate Gaussian distribution. It follows that the conditional and marginal distributions are also Gaussian, making it possible and sufficient to estimate their means and covariance (Rue and Held, 2005). The conditional state and observations densities respectively are:

$$S_t \mid S_{t-1} \sim MVN\ (AS_{t-1}, BWB^{'}) \text{ and } Y_t \mid S_t \sim MVN\ (GS_t,\ HVH^{'})$$

Let the posterior assumed to be Gaussian with mean and covariance at time t-1 be denoted respectively as $\bar{S}_{t-1}$ and P_{t-1}. Using standard results, we can obtain the one-step predictive density $\pi(S_t|Y_{1:t-1}) = \int \pi(S_t|S_{t-1})\pi(S_{t-1}|Y_{1:t-1})\, dS_{t-1}$ which can be evaluated directly as Gaussian with mean $\bar{S}_{t|t-1} = A\bar{S}_{t-1}$ and covariance $P_{t|t-1} = AP_{t-1}A^{'} + BWB^{'}$.

The computation of the filtering density which is proportional to $\pi(Y_t|S_t)\pi(S_t|Y_{1:t-1})$ can be obtained through the joint density of Y_t and S_t conditional on S_{t-1} and the result is multivariate normal (Rue and Held, 2005):

$$\begin{pmatrix} S_t \\ Y_t \end{pmatrix} \mid S_{t-1} \sim MVN\left[\begin{pmatrix} \bar{S}_{t|t-1} \\ G\bar{S}_{t|t-1} \end{pmatrix}, \begin{pmatrix} P_{t|t-1} & GP^{'}_{t|t-1} \\ P_{t|t-1}G^{'} & K_t \end{pmatrix}\right],$$

where the matrix $K_t = GP_{t|t-1}G^{'} + HVH^{'}$ is the so-called the Kalman gain.

The resultant posterior mean and covariance (Rue and Held, 2005; Taylor, 2010) are respectively

$\bar{S}_t = \bar{S}_{t|t-1} + P_{t|t-1}G^{'}K_t^{-1}(Y_t - G\bar{S}_{t|t-1})$ and

$P_t = P_{t|t-1} - P_{t|t-1}G^{'}K_t^{-1}GP_{t|t-1}$

Time variant extensions such as A_t and G_t respectively for A and G allows time varying disturbances as well as system and observation parameters. Detailed discussion and reviews on Kalman filtering are published elsewhere (Harvey, 1990; Kalman and Bucy, 1961).

The Kalman filter can be extended to model state space problems through the use of extended Kalman filter and unscented Kalman filter when the assumptions of linear and Gaussian state and observation equations are violated. Detailed discussion and reviews are presented elsewhere (Anderson and Moore, 1979; Ito and Xiong, 2000; Jazwinski, 1970).

3.6 Conclusion

There are 4 papers in this thesis. The first paper presented in chapter 4 is based on (3.1). Paper 2 which is presented in chapter 5 is based on (3.4) while paper 3 presented in chapter 6 assumed the model in (3.6). Paper 4 in Chapter 7 is based on (3.3).

Reference

Anderson, B. D. O., and J. B. Moore (1979). Optimal Filtering, Englewood Cliffs, New Jersey, Prentice Hall inc.

Cressie, N.(1990). The origins of kriging: Mathematical Geology, v. 22, p. 239-252.

Cressie, N. A. (1993). Statistics for Spatial Data: New York, John Wiley & Sons.

Davis, J. A., J. L. Spaeth, and C. Huson (1961). A technique for analyzing the effects of group composition, American Sociological Review, p. 215-225.

Diggle, P., P. J. Ribeiro, and MyiLibrary. (2007). Model-based geostatistics, Springer series in statistics, New York, NY, Springer.

Goldstein, H. (1987), Multilevel models in educational and social research., London, Griffin.

Goldstein, H. (2003). Multilevel Statistical Models: London, Arnold.

Harvey, A. C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press.

Ito, K., and K. Xiong (2000). Gaussian filters for nonlinear filtering problems., IEEE Transactions on Automatic Control p. 910-927.

Jazwinski, A. H. (1970). Stochastic processes and filtering theory: New York, Academic Press.

Kalman, R., and R. Bucy (1961). New results in linear filtering and prediction theory Journal of Basic Engineering,Transactions ASME Series D, p. 95-108.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems Journal of Basic Engineering,Transactions ASME Series D, p. 35-45.

KNOX, E. (1964). THE DETECTION OF SPACE-TIME INTERACTIONS: The Royal Statistical Society Series C-Applied Statistics, v. 13, p. 25-29.

Longford, N. T. (1993). Random Coefficient Models: New York, Oxford University Press.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals, American Sociological Review, p. 351-357.

Rue, H., and L. Held (2005). Gaussian Markov Random Fields, Chapman & Hall.

Snijders, T. A. B., and R. J. Bosker (2012). Multilevel analysis: an introduction to basic and advanced multilevel modeling: London, SAGE.

Taylor, B. M. (2010). Sequential Methodology and Applications in Sports-Rating. PhD Thesis, Lancaster University.

Thum, Y. M. (1997). Hierarchical Linear Models for Multivariate Outcomes, Journal of Educational and Behavioral Statistics, p. 77-108.

Chapter 4 Paper 1: Childhood Malnutrition and its Determinants among Under-Five Children in Ghana

This chapter is based on the following paper:

Aheto JMK, Keegan TJ, Taylor BM, Diggle PJ. (2015). Childhood Malnutrition and Its Determinants among Under-Five Children in Ghana. *Paediatric and Perinatal Epidemiology*; **29**:552-561.

Abstract

Background: Childhood malnutrition adversely affects short and long term health and economic wellbeing of children. Malnutrition is a global challenge and accounts for around 40% of under-five mortality in Ghana. Limited studies are available indicating determinants of malnutrition among children. This study investigates prevalence and determinants of malnutrition among children under-five with the aim of providing advice to policy makers and other stakeholders responsible for the health and nutrition of children.

Methods: The study used data from the 2008 Ghana Demographic and Health Survey (GDHS). Analyses were conducted on 2,083 children under-five years nested within 1,641 households with eligible anthropometric measurements, using multilevel regression analysis. Results from the multilevel models were used to compute probabilities of malnutrition.

Results: This study observed that 588 (28%), 276 (13%) and 176 (8%) of the children were stunted, underweight and wasted respectively. Older ages are associated with increased risk of stunting and underweight. Longer breast feeding duration, multiple births, experience of diarrhoeal episodes, small size at birth, absence of toilet facilities in households, poor households, and mothers who are not covered by national health insurance are associated with increased risk of malnutrition. Increase in mother's years of education and body mass index are associated with decreased malnutrition. Strong residual household-level variations in childhood nutritional outcomes were found.

Conclusion: Policies and intervention strategies aimed at improving childhood nutrition and health should address the risk factors identified and the need to search for additional risk factors that might account for the unexplained household-level variations.

Keywords

Epidemiology, Childhood malnutrition, Under-five children, Developing countries, Nutritional status, Multilevel modelling, Malnutrition determinants, Public health.

Funding

Funded by Faculty of Health and Medicine, Lancaster University, United Kingdom. Except the funding of the study, the funder did not play any other role in this study including the published article.

4.1 Introduction

Childhood malnutrition is a major public health issue. Not only does malnutrition bring with it illness and development issues for the children who suffer it but, along with poor sanitation and diseases such as malaria, it is an important cause of childhood mortality. Globally, about 30% of deaths among under-five children are attributable to malnutrition (Black et al., 2008). This percentage is higher in developing countries, at around 50% (WHO, 2005).   In Ghana, malnutrition accounts for around 40% of deaths in under-fives (Antwi, 2008; ORC Macro, 2005; UNICEF-Ghana, 2008).

Childhood malnutrition is the result of multiple factors. Environmental conditions, socio-economic circumstances and feeding practices are all important factors (Bomela, 2009; Kandala et al., 2011). The consequences of malnutrition in infants and young children are to limit and set back physical growth and neurological development, to lower intelligence quotient (IQ) and to increase susceptibility to disease (de Onis and Blössner, 2003; Okoromah et al., 2011; The World Bank, 2006, 2014; UNICEF et al., 2012; UNICEF., 2006).

In Ghana, child mortality rates are stabilising (Food and Agriculture Organization (FAO), 2009), but malnutrition in children under 5 is still one of the major causes of death in children. While information on the prevalence of malnutrition in the country, and more widely the region, are known (Antwi, 2008; Ghana Statistical Service et al., 2009) less has been published about the social, economic and environmental factors facing the residents of Ghana that may explain the country's relatively high rates of childhood malnutrition and, consequently, childhood and maternal mortality.

In this paper we aim to determine the major risk factors for childhood malnutrition in Ghana; particularly those individual and household characteristics indicated as risk factors for being malnourished under the age of 5 years.

The data we have used comes from the 2008 Ghana Demographic and Health Survey, which provide information on both children and the households in which they live. A specific aim of the paper is to apply multilevel modelling techniques to explore risk factors for malnutrition so that we can understand the impacts on malnutrition of children, both from factors that are specific to an individual, such as age, while recognising that some factors that affect individuals such as the amount of food available, are common to all members of a household.

4.2 Methods

4.2.1 Study population

This is a population-based cross-sectional study using data from the 2008 Ghana Demographic and Health Survey (GDHS) (Ghana Statistical Service et al., 2009). The data were provided by the Ghana Statistical Service and contain information on maternal and child health, nutritional status of women and children, breastfeeding practices, fertility preferences, awareness and use of family planning methods, childhood mortality and domestic violence.

The GDHS survey methods are published elsewhere (Ghana Statistical Service et al., 2009). To summarise, representative samples of 12,323 households from 411 sampling units (communities) were selected nationwide. From these, 11,778 eligible households were interviewed, amongst which 5,175 and 6,603 were in urban and rural communities, respectively. Data were collected from 4,916 women aged 15-49 and

4,568 men aged 15-59. Data on 2,992 children from 1,000 and 1,992 urban and rural communities, respectively, were then generated from the sampled women during the main survey (Ghana Statistical Service et al., 2009).

756 children were excluded because of missing exposure data. A further 153 were excluded because their Z-scores were outside the range of biologically plausible values as determined by the World Health Organization (WHO) (WHO, 2010).

The analysis reported here was therefore based on 2,083 under-five children residing in 1,641 households and 400 communities across Ghana. Of these 1,641 households, 416 (25.4%) contained two or more under-five children.

4.2.2 Outcome variables

The primary outcome of interest was malnutrition among children under-five years of age. Malnutrition was determined using gender-specific Z-scores to obtain three WHO-derived indicators: height-for-age (HAZ), weight-for-age (WAZ) and weight-for-height (WHZ) Z-scores (WHO Multicentre Growth Reference Study Group, 2006). The Z-scores was calculated using the macro provided by WHO (WHO, 2010). These measures were available for all 2,083 children in the final analysis.

The scales height-for-age (HAZ), weight-for-age (WAZ) and weight-for-height (WHZ) are also referred to as stunting, underweight and wasting, respectively. A child is categorised as malnourished on any of these scales if his/her Z-score is below minus two standard deviations from the median of the reference population (WHO Multicentre Growth Reference Study Group, 2006).

4.2.3 Explanatory variables

We included as potential explanatory variables those risk factors identified in the literature as significant predictors of under-five nutritional status in developing countries (Adekanmbi et al., 2013; Alom et al., 2012; Babatunde et al., 2011; Bomela, 2009; Kabubo-Mariara et al., 2009; Kandala et al., 2011) together with additional variables that we considered to be related to the nutritional status and health of children, such as the national health insurance scheme (NHIS) status of the mother. Mothers who are covered by NHIS will be more likely to visit health facilities, thereby obtaining prescriptions and better health care when they are ill. We expect this to result in better nutrition and health outcomes for mothers and their children. To arrive at the final set of included risk factors in our multilevel model, we used a backward elimination method. Presented in Supporting Information Appendix S1 is the description of explanatory variables used in this analysis.

4.2.4 Statistical analysis

We analysed the outcomes of interest to examine whether there were differences in the nutritional outcome of children by both individual and household level risk factors and whether or not children from different households exhibit different nutritional outcomes. To do this, we applied a random intercept multilevel regression model (Goldstein, 2003) presented in Chapter 3 of this thesis to analyse our outcome variables on a continuous scale and obtained the parameter estimates in our model using maximum likelihood. Among competing covariance structures, the random intercept model with a within-group variance component provided a good fit to our data. Details of the model are shown in Equation 4.1 in the Supporting Information Appendix. The choice of multilevel regression analysis in this study is appropriate because we have data on individual children under-five nested within households. Not recognising the hierarchical structure of the data could lead to underestimation of the standard errors for the regression coefficients, which in turn would lead to spurious statistical significance and incorrect inference. We presented the description of the summary measure (Steele et al., 2007) of variance explained by households in WAZ, HAZ or WHZ among children in the Supporting Information Appendix S2 - Equation (4.4). We computed the probabilities of malnutrition among children as functions of their risk factors from the multilevel model (see Supporting Information Appendix S2 in Equation (4.3)). To examine the relationship between the probabilities of malnutrition (any of stunting, underweight and wasting) among children and the individual risk factors, the computed probabilities were plotted against each of the statistically significant risk factors for each of the three outcomes.

The statistical analysis was carried out using R (R Core Team, 2014). For fitting the random intercept multilevel model, we used the R package 'nlme'. We used likelihood ratio tests (LRTs) to assess the significance of risk factor effects. To assess the significance of the household-level random effect, we divided the p-values from the LRT by two because the null hypothesis that $\sigma^2_H = 0$ is on the boundary of the parameter space (Self and Liang, 1987). For a general review of the statistical methods used in this analysis, see Chapter 3 section 3.1.1 of this thesis.

4.3 Results

4.3.1 Sample characteristics

In the data-set, 588 (28.2%), 276 (13.3%) and 176 (8.4%) of the children were moderately stunted, moderately underweight and moderately wasted, respectively (see Table 4.1) based on the WHO classification for assessing severity of malnutrition by prevalence ranges among children under-five (de Onis and Blössner, 1997). A total of 784 (38%) children had mothers with no formal education.

Table 4.1 Percentage distribution of children under-five by selected background characteristics (n=2083)

Characteristics	n (%)	Characteristics	n (%)
Individual child level		Household wealth status	
Sex		Average/rich households	934 (44.8)
Male	1,045 (50.2)	Poor households	1,149 (55.2)
Female	1,038 (49.8)		
		Piped source of drinking water	666 (32.0)
Age (months)		Non-piped source of drinking water	1,417 (68.0)
<24	843 (40.5)		
24 or more	1,240 (59.5)	Flush/pit toilets facility in household	1,355 (65.0)
		No toilet facility in household	728 (35.0)
Still breast-feeding at the time of survey	801 (38.5)	Mothers without health insurance	1,232 (59.1)
Not breast-feeding	1,282 (61.5)	Mothers with health insurance	851 (40.9)
Type of birth			
Single birth	2,015 (96.7)	Mothers' BMI	
Multiple birth	68 (3.3)	BMI<18.5 kg/m^2 (underweight)	154 (7.4)
		BMI≥18.5&<25 kg/m^2 (normal)	1,412 (67.8)
		BMI≥25.0 kg/m^2(overweight)	515 (24.8)
Had diarrhoea	456 (21.9)		
No diarrhoea	1,627 (78.1)	Community characteristic	
		Rural	1,391 (66.8)
Number of dead children		Urban	692 (33.2)
One or more	498 (23.9)		
None	1,585 (76.1)		
Had fever	444 (21.3)		
No fever	1,639 (78.7)		
Small size at birth	304 (14.6)		
Large/average size at birth	1,779 (85.4)		
Delivered at health care facility	1,110 (53.3)		
Not delivered at health care facility	973 (46.7)		
Stunted (HAZ below minus 2 SD)	588 (28.2)		
Underweight (WAZ below minus 2 SD)	276 (13.3)		
Wasted (WHZ below minus 2 SD)	176 (8.4)		
Maternal or household level			
Maternal education			
Secondary or higher	805 (38.7)		
Primary	494 (23.7)		
No formal education	784 (37.6)		

SD = standard deviations

All 2,083 children considered in the analysis received some amount of breast feeding, whilst 1,854 (89%) were breastfed for more than 6 months, amongst whom 1,038 (56%) belonged to poor households. The number of children who had experienced diarrhoea and fever episodes within two weeks of the survey were 456 (22%) and 444 (21%)

respectively. A total of 973 (47%) children were not delivered at health care facilities (see Table 4.1).

4.3.2 Risks factors associated with WAZ, HAZ and WHZ for multilevel models

Table 4.2 presents results of the multilevel modelling for the three measures of malnutrition WAZ (underweight), HAZ (stunting) and WHZ (wasting). Statistically significant risk factors for WAZ (underweight) were: older age; longer breast-feeding duration; multiple births; diarrhoea history; and small size at birth. Mother's BMI and years of education were positively associated with WAZ. A history of fever, no toilet facility in household, mothers without health insurance cover and mother's current age were not associated with a child's probabilities of being underweight.

Risk factors negatively associated with HAZ were: older ages of children; longer breast feeding; multiple births; small size at birth; poor household and mothers without health insurance cover. Mother's BMI and current age were positively associated with HAZ.

Factors associated with WHZ (wasting) were: a history of diarrhoea; small size at birth; and no toilet facility in household. Older ages of children and mother's BMI were also positively associated with WHZ. Longer breast-feeding, multiple births, fever history and mother's current age were not associated with wasting.

The results show significant residual household-level variation. Table 4.3 shows that 32%, 23%, and 20% of the variance in the nutritional outcome of children for WAZ, HAZ and WHZ respectively could be attributed to the residual household-level variations after adjusting for child and household covariates.

Table 4.2 The effect estimate (β) for the associations between risk factors and child nutritional status (weight-for-age or WAZ (a measure from which underweight is derived), height-for-age or HAZ (a measure from which stunting is derived) and weight-for-height or WHZ (a measure from which wasting is derived)) among under-five children in Ghana using random intercept multilevel regression models (n=2083).

	Parameter estimates		
Risk factors	WAZ (95% CI)	HAZ (95% CI)	WHZ (95% CI)
Child level			
Age in years	-0.041 (-0.077, -0.005)	-0.169 (-0.217, -0.121)	0.125 (0.082, 0.168)
Months of breast feeding	-.015 (-0.021, -0.008)	-0.043 (-0.052, -0.034)	0.007 (0.000, 0.015)
Type of birth			
Single birth	1.000 (Reference)	1.000 (Reference)	1.000 (Reference)
Multiple birth	-0.487 (-0.747, -0.227)	-0.470 (-0.811, -0.130)	-0.266 (-0.561, 0.028)
Diarrhoea			
No	1.000 (Reference)	-	1.000 (Reference)
Yes	-0.148 (-0.253, -0.044)	-	-0.163 (-0.285, -0.041)
Fever			
No	1.000 (Reference)	-	1.000 (Reference)
Yes	-0.094 (-0.199, 0.010)	-	-0.093(-0.215, 0.029)
Size of child at birth			
Large/average	1.000 (Reference)	1.000 (Reference)	1.000 (Reference)
Small	-0.310 (-0.429,-0.191)	-0.219 (-0.378, -0.060)	-0.268 (-0.407, -0.129)
Maternal/ household			
Mother's body mass index	0.043 (0.032, 0.054)	0.034 (0.019, 0.049)	0.031 (0.018, 0.043)
Mothers' education (years)	0.014 (0.002, 0.025)	-	-
Household wealth status			
Rich or average	-	1.000 (Reference)	-
Poor	-	-0.164 (-0.291, -0.037)	-
Mothers' current age	0.005 (-0.002, 0.012)	0.021 (0.012, 0.030)	-0.007(-0.014, 0.001)
Type of toilet facility			
Flush or pit	1.000 (Reference)	-	1.000 (Reference)
No facility	- 0.099 (-0.203, 0.005)	-	-0.162 (-0.272, -0.052)
Number of dead siblings			
None	1.000 (Reference)	-	-
One or more	0.097 (-0.015, 0.209)	-	-
Mother has health insurance			
Yes	1.000 (Reference)	1.000 (Reference)	-
No	-0.090 (-0.183, 0.003)	-0.176 (-0.297, -0.055)	-

4.3.3 Probabilities of malnutrition for significant risk factors

We plotted the probabilities of stunting, underweight and wasting against the risk factors identified as statistically significant in the multilevel models presented in Table 4.2 (see Figures 4.1 and 4.2). The probabilities of the three measures of malnutrition all decreased with increasing mother's BMI. Probabilities of stunting increased with child's age, conversely probabilities of wasting decreased with child's age. Increasing duration of breastfeeding was associated with an increase in risk of underweight and stunting. The risk of a child being underweight decreased with higher levels of mother's education but this was not seen when malnutrition was measured as stunting or wasting. Mother's age was a risk factor only for stunting.

The risk of underweight and stunting is higher among children who were products of multiple births compared to those who were singletons. Children who had diarrhoea were more at risk of underweight and wasting. The probabilities of underweight, stunting and wasting were higher among children born small in size at birth compared to those born average or large in size. Children from poor homes were more at risk of stunting compared to those from average or rich homes and children from mothers who were not covered by national health insurance had higher risk of stunting compared to those from mothers who had national health insurance. The probabilities of wasting increased among children from homes with no toilet facilities compared to those from homes with pit or flush toilet facilities.

Table 4.3 Variance estimates with 95% confidence intervals for the random intercept multilevel regression models of weight-for-age or WAZ (a measure from which underweight is derived) , height-for-age or HAZ (a measure from which stunting is derived), and weight-for-height or WHZ (a measure from which wasting is derived) (n=2083).

	Variances		Intra-household correlation coefficients
Outcomes	Child (95% CI)	Household (95% CI)	Explained variation (%)[a]
WAZ	0.65 (0.58, 0.73)	0.31 (0.24, 0.40)	32
HAZ	1.31 (1.16, 1.47)	0.39 (0.26, 0.58)	23
WHZ	1.03 (0.91, 1.16)	0.25 (0.16, 0.40)	20

[a]Ratio of household-level variance to total variance multiplied by 100.

4.4 Discussion

We set out to investigate the determinants of childhood malnutrition in young children in Ghana. Our data on both individuals and their households allowed us to investigate both individual risk factors and those that stem from shared exposures by household. We also examined household-level random effects, which represent variation in household-level outcomes that cannot be explained by the available household-level covariates.

We found that out of the 2,083 children considered in the analysis, 588 (28%), 276 (13%) and 176 (8%) were stunted, underweight and wasted, respectively; suggesting moderate form of malnutrition for each of these outcomes among under-five population of Ghanaian children based on the WHO classification for assessing severity of malnutrition by prevalence ranges among children under 5 years of age (de Onis and Blössner, 1997). This evidence suggests that childhood malnutrition still remains a serious public health challenge in Ghana.

Individual child risk factors that were predictive of malnutrition were: child’s age; type of birth; child’s experience of diarrhoeal episodes; size of child at birth; and months of breast feeding. Household-level variables that were associated with malnutrition were:

mother's education, current age, BMI and national health insurance status; household toilet facility ownership and wealth status.

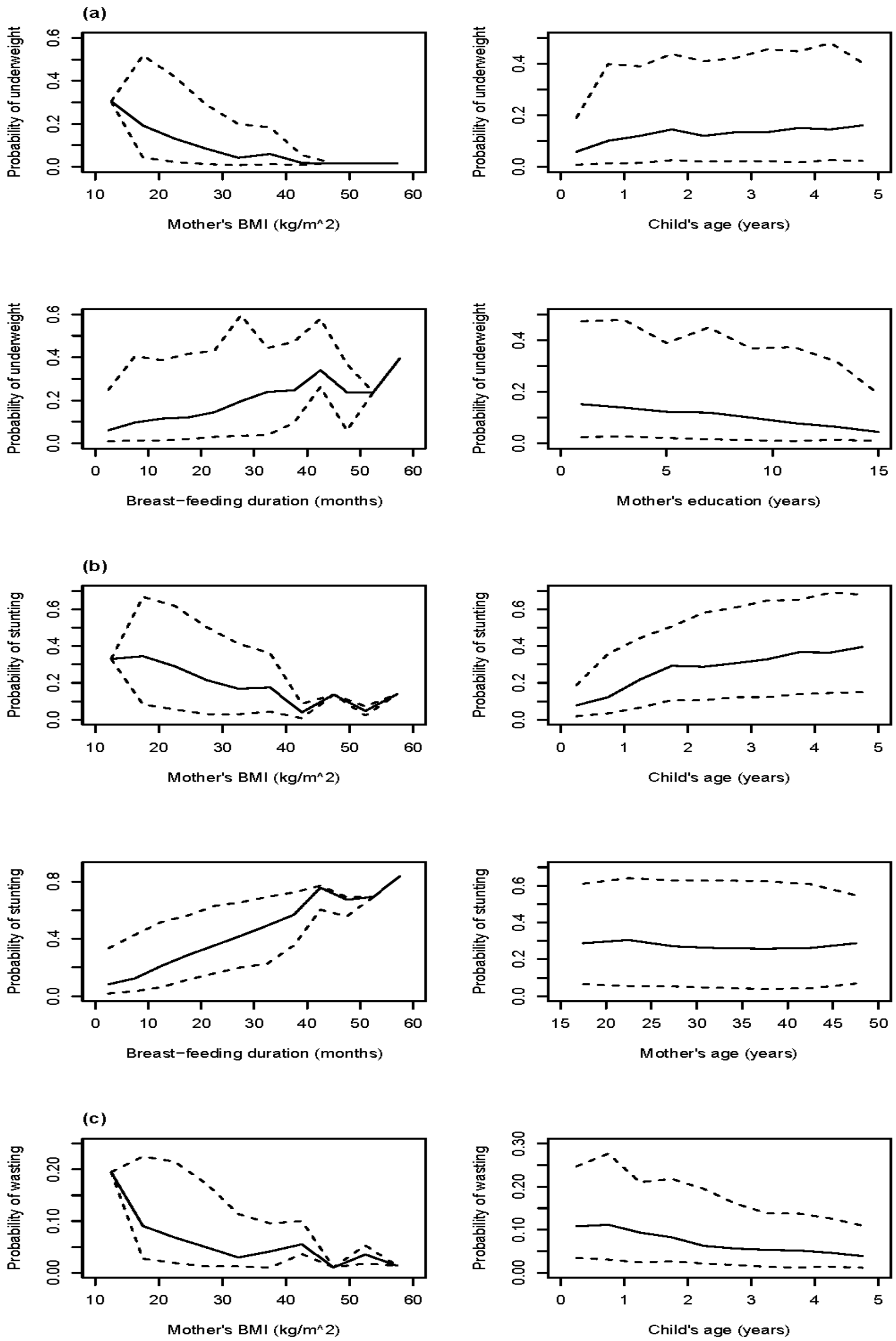


Figure 4.1 Plots showing probabilities of malnutrition for continuous risk factors, by malnutrition measure: a) Underweight (WAZ), b) Stunting (HAZ), c) Wasting (WHZ).
The solid and dotted lines respectively respresent the median values for the risk factors and their associated 95% confidence intervals.

We were particularly interested in the household-level variation in childhood malnutrition. Households constitute key determinants of socioeconomic disparities in health and general wellbeing of children as they influence each child's opportunities and, to some extent, govern exposure to risks and resources over the life course (Adekanmbi et al., 2013; Kandala et al., 2011). This study has shown significant residual household-level variation in childhood nutrition, implying that nutritional outcomes of children vary across households in Ghana after adjusting for child and household characteristics. Our analysis shows that 32%, 23%, and 20% of the variation in the nutritional outcome of children for WAZ, HAZ and WHZ, respectively could be attributed to unobserved household-level factors. These could be social or environmental or both. For instance, the factors could be related to the location of the household and may indicate geographical differences in factors related to malnutrition in children.

The strengths of our study are that it is a large, population-based study with national coverage and good quality data on a number of child and household characteristics. The method we used, multilevel modelling, allowed the identification of household-level variation from presently unidentified factors.

The limitations of the study include the fact that, and it is known that wealth status is associated with malnutrition (Adekanmbi et al., 2013; Bomela, 2009), it is difficult to directly measure household wealth status in Ghana. Because of this we used an asset-based index, which is generally considered a good proxy for household wealth status in developing countries. In addition we did not have complete data for all children for all variables, for example only 1,163 out of the 2,992 children in the main survey have measurements on their birthweight.

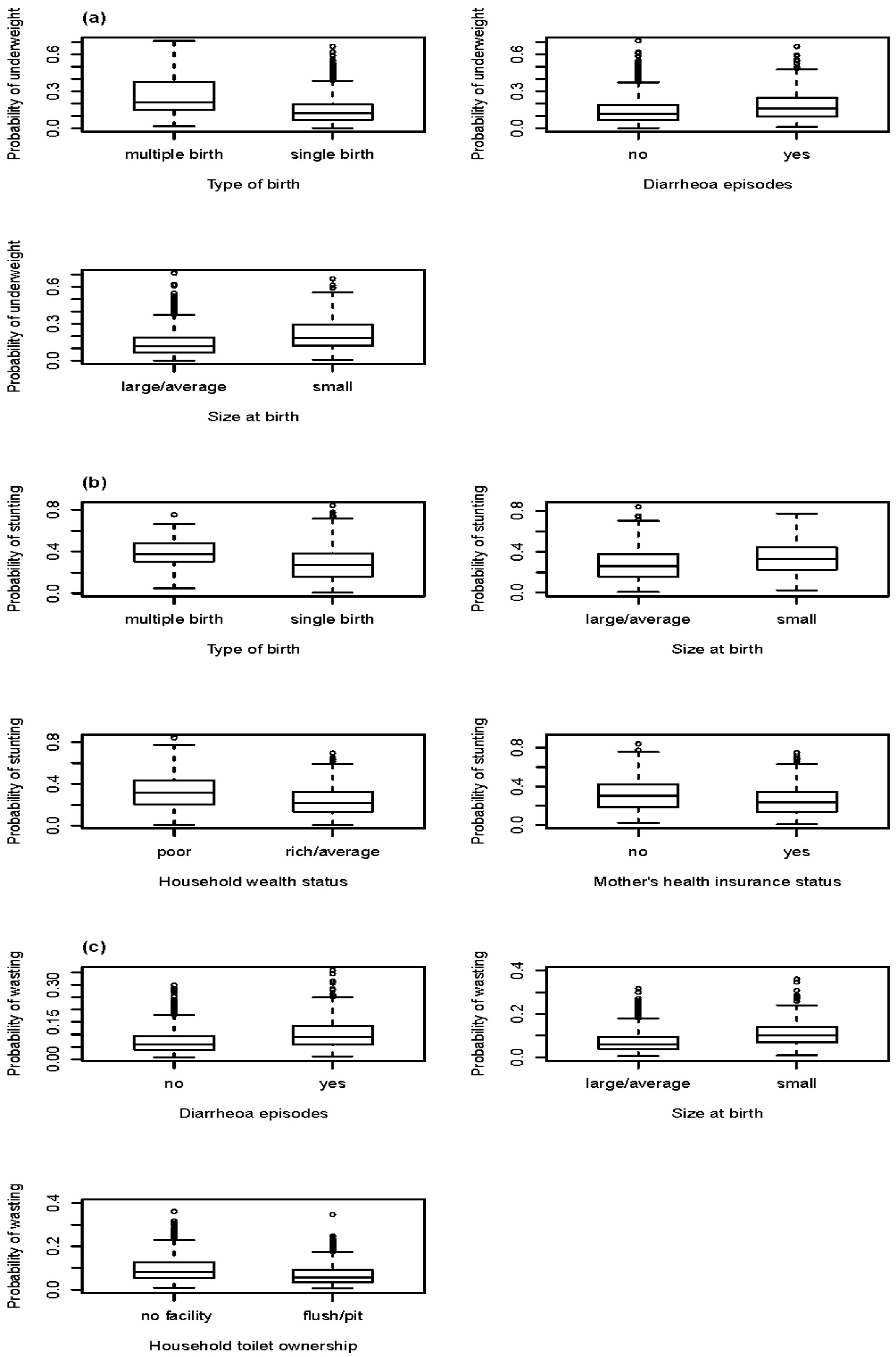


Figure 4.2 Box plots showing probabilities of malnutrition for categorical risk factors, by malnutrition measure: a) Underweight (WAZ), b) Stunting (HAZ), c) Wasting (WHZ).

To maximise the amount of data available for our analysis of birthweight, we used size of child at birth as perceived and reported by the mother as a proxy for birthweight. This variable is self-reported and could introduce reporting bias. We conducted a logistic regression analysis using size of child (coded as large/average=1, small=0 for the logistic regression analysis) as outcome and actual birth weight (kg) as covariate and found a strong, albeit imperfect, association. The estimated log-odds of large/average vs. small increased by 1.96 for every 1kg increase in actual birthweight. Expressed more tangibly, the probability of large/average self-reported birthweight increased from 0.14 at a birthweight of 1kg to 0.89 at a birthweight of 3kg (see Supporting Information Appendix S3).

Our study broadly supports those of previous research on malnutrition in developing countries. For example, children who were products of multiple births experienced more diarrhoeal episodes, and were smaller in size at birth, whilst children who were breast fed for longer duration had higher risks of malnutrition (Adekanmbi et al., 2013; Griffiths et al., 2004; Hong, 2006; Kabubo-Mariara et al., 2009; Kandala et al., 2011; Madise and Mpoma, 1997; Mbago and Namfua, 1992; Uthman, 2007). The increase in risks of malnutrition among children who were products of multiple births could be the result of either low birth weight or competition for nutritional intake, which happens more among children who are products of multiple births than those of singletons (Adekanmbi et al., 2013). The increase in risks of malnutrition observed among children who experienced diarrhoeal episodes could be due to the fact that diarrhoea normally results in wastage of food nutrients and loss of appetite.

The UN Food and Agriculture Organisation has reported that while breastfeeding in Ghana is common practice, only half of children under 6 months are exclusively breastfed, and complementary feeding practices are inadequate (Food and Agriculture

Organization (FAO), 2009). Our finding that probabilities of underweight and stunting increased among children who were breast fed for longer duration could be the result of poverty among households in Ghana; consequently, mothers may continue to breast feed beyond the recommended 6 months without supplementation (Van de Poel et al., 2008). In contrast, another study conducted in Jamaica reported a positive association between longer duration of breast feeding and child nutritional status (Melville et al., 1988). This could suggest a difference in results between low-income and middle-income countries.

Except for wasting, our study observed that children's age is associated with higher malnutrition. The positive association of stunting (low height for age) with age before two years shows that children are not growing satisfactorily after cessation of breastfeeding. This could be the result of a deficit in proper complementary food and presence of progressive childhood diseases (Adekanmbi et al., 2013; Alom et al., 2012).

Babies small in size at birth were more likely to be malnourished. We looked at the the effect of the actual birthweight on the probability of being born large/average versus small in size at birth as reported by mother (using the children for whom we did have a birthweight) and found that the probability of a child being born large/average in size at birth increases with increasing actual birthweight (see Supporting Information Appendix S3). We also found that decreasing birthweight was associated with increased risk of malnutrition. This gives some justification for size at birth as a proxy for birthweight, and is consistent with previous findings that low birthweight is a risk factor for malnutrition in under-five children (Adekanmbi et al., 2013; Madise and Mpoma, 1997).

Children from mothers with high levels of education, older ages and higher BMI had decreased risks of malnutrition, while children from poor homes, with no toilet facilities and whose mothers have no national health insurance cover had increased risks of malnutrition.

We were curious as to whether diarrhoea and toilet facilities are independent risk factors for malnutrition, or diarrhoea is an intermediate effect of lacking a toilet facility at home. In our analysis the association (results not shown) between diarrhoea and malnutrition was reduced but remained significantly raised when adjusted for toilet facilities at home. Toilet facilities in households have been documented in the literature as contributing to improved nutritional outcome of children and serving as a proxy for high socio-economic status in developing countries (Madise and Mpoma, 1997).

The decrease in risks of malnutrition for children from households with mothers who were covered by the NHIS is likely to be because mothers who are covered are more likely to visit health facilities and to seek health care for them and their families, which should result in better health outcomes for mothers and their children (Blanchet et al., 2012; Mensah et al., 2010). The increase in risk of malnutrition in children from poor households is likely to be because children from these households are more likely to have low quality and insufficient food intake, poorer living conditions, greater exposure to diseases and inadequate or complete lack of access to basic health services (Adekanmbi et al., 2013; de Onis and Blössner, 2003; The World Bank, 2006; Uthman, 2007).

A decrease in malnutrition among children from mothers with high level of education suggests that improving mother’s education will improve the level of child nutritional outcomes (Adekanmbi et al., 2013; Babatunde et al., 2011; Griffiths et al., 2004; Kabubo-Mariara et al., 2009; Kandala et al., 2011). It has been shown that improvements in women’s education bring many advantages to their lives and to society more generally (Adekanmbi et al., 2013; Madise and Mpoma, 1997). We also found that the risk of malnutrition increased as mothers' BMI decreased, especially once BMI was below approximately 20 kg/m^2. Maintaining improvement in education and adequate nutrition in mothers is important for preventing malnutrition in children (The World Bank, 2014).

The findings from this study have key policy and intervention implications for improving childhood nutrition and health in Ghana and more widely. Public health measures should build on the work already done to implement WHO recommendations on exclusive breast feeding up to 6 months of age. Also, the need to feed infants with nutritionally adequate and safe complementary foods from six months of age together with continued breastfeeding up to age two or beyond must be emphasized. Additionally, countries should strive to provide free health care services for pregnant women, mothers and children under five years of age.

Addressing poverty is one important area that governments and non-governmental organizations (NGOs) responsible for health and nutrition of children should consider. Many of the environmental and social determinants of childhood malnutrition observed here are addressed by the UN Millennium Goals, e.g. eradication of poverty and hunger, gender equality and reduction of childhood mortality. Beyond economic redistribution, these could include subsidies for children's food and assistance for households in installing sanitary toilets in their homes to help improve the health and nutrition of young children. Better toilet facilities in homes will prevent the spread of infectious diseases, which negatively affect the health and nutrition of children. To increase levels of women's education more generally, governments such as Ghana's could make basic and secondary education more accessible, compulsory and affordable.

Our results show that there are unanswered questions about the reasons for variation in malnutrition associated with the house in which a child lives. We are undertaking research aimed at investigating the geographical predictors of malnutrition in Ghana. Results from that research may contribute useful information to the discussion about household effects. More research on household factors not measured in this study and their possible effects on malnutrition may also be warranted.

We believe that the relationship between extended breastfeeding and increased likelihood of malnutrition is not causal - that is it confounded by an absence of complimentary foods rather than an adequacy of breast feeding and further research is warranted in this area to explore this issue. This finding should therefore be interpreted with greatest caution.

4.5 Conclusion

Policies and intervention strategies by policymakers that are aimed at improving nutrition and health of children should address the risk factors identified in this study. There is also a need to identify as-yet unidentified risk factors that might account for the unexplained household-level variations in childhood nutritional outcomes.

4.6 Appendix

4.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis.

Table 4.4 Description of explanatory variables used in this analysis

Child characteristics
Age: Age in years and is continuous
Breastfeeding duration: breastfeeding duration in months and continuous
Type of birth: 1= multiple birth and 0= single birth
Diarrhoea episodes: 1= child had diarrhoea and 0= child did not have diarrhoea
Fever episodes: 1= child had fever and 0= child did not have fever
Size at birth: 1= small and 0= average/large
Maternal/household characteristics
Mother's education: mother's education in years and is continuous
Mother's body mass index (BMI): mother's BMI and is continuous
Mother's age: mother's age in years and is continuous
Mother's health insurance status: 1= mother did not have health insurance and 0= mother had health insurance
Household wealth status: 1= poor and 0= average/rich
Type of toilet facility in household: 1= no facility and 0= flush/pit facility
Number of dead siblings in household: 1= one or more and 0= none

4.6.2 Online Supporting Information Appendix S2: Supplementary material related to the random intercept multilevel regression model and computation of probabilities of malnutrition

Our model is

$$Y_{ij} = x_{ij}^{'} \beta + H_j + \varepsilon_{ij} \qquad (4.1)$$

with $H_j \sim N(0, \sigma^2_H)$, and $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$. Y_{ij} is the nutritional status on a continuous scale (WAZ, HAZ or WHZ) for child i in household j, β is a vector of regression coefficients, x_{ij} is a vector of risk factors at the child or household level, H_j is a random effect for the jth household with associated variance σ^2_H and ε_{ij} is an individual-level residual with associated variance σ^2_ε. In other words, H_j is the deviation from the overall mean for the jth household and ε_{ij} is the deviation of the ith child from the mean of the jth household. The quantities σ^2_H and σ^2_ε were used to obtain the intra-class correlation coefficient which is a useful tool for assessing the degree of homogeneity within classes such as households in this study.

We used maximum likelihood to obtain estimates $\hat{\beta}$, $\hat{\sigma}_H^2$, $\hat{\sigma}_\varepsilon^2$ of the parameters in (4.1), and computed the minimum mean square error predictors $\widehat{H}_j$ for each household j as the expectations of H_j conditional on the data and maximum likelihood parameter estimates. We are interested in predicting the nutritional status for a child *i* belonging to household *j* with the same covariate attributes as child *i* from that household i.e. x_{ij}. This ideally would involve computing $[Y_{ij} \mid \text{data}]$, where the notation [.] means "the distribution of". This distribution is analytically intractable. However, $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_H^2$ are well estimated from the data, and we therefore use a "plug-in" approximation $[Y_{ij} \mid \text{data}, \hat{\sigma}_\varepsilon^2, \hat{\sigma}_H^2]$; in doing this, we are still able to take into account the uncertainty in the estimates of β and the predictors of the H_j. This is achieved as follows.

Our estimates of β and H_j are asymptotically Normal with means $\hat{\beta}$ and $\widehat{H}_j$ and variance equal to the inverse of the Fisher information. We replace the Fisher information by the observed information, denoted here by v_β and $\sigma_{H_j}^2$, and use the approximations $[\beta|\text{data}] \sim N(\hat{\beta}, v_\beta)$ and $[H_j|\text{data}] \sim N(\widehat{H}_j, \sigma_{H_j}^2)$. Using these approximations, the required distribution, $[Y_{ij} \mid \text{data}, \hat{\sigma}_\varepsilon^2, \hat{\sigma}_H^2]$, becomes

$$[Y_{ij} \mid data, \hat{\sigma}_H^2, \hat{\sigma}_\varepsilon^2] = \iint [Y_{ij}, H_j, \beta|\text{data}, \hat{\sigma}_H^2, \hat{\sigma}_\varepsilon^2]\, dH_j\, d\beta$$

$$= \iint [Y_{ij} \mid H_j, \beta, \hat{\sigma}_H^2, \hat{\sigma}_\varepsilon^2, data]\, [\beta, \widehat{H}_j \,|\text{data}, \hat{\sigma}_H^2, \hat{\sigma}_\varepsilon^2]\, dH_j\, d\beta$$

$$= \iint [Y_{ij} \mid H_j, \beta, \hat{\sigma}_\varepsilon^2, \text{data}]\ [H_j \,|\text{data}\,][\beta|\text{data}]\, dH_j\, d\beta \qquad (4.2)$$

The integrand in (4.2) invokes the further approximation of $[H_j \,|\text{data}, \hat{\sigma}_H^2, \hat{\sigma}_\varepsilon^2\,]$ by $[H_j \,|\text{data}\,]$. Re-writing (4.2) gives

$\iint N(Y_{ij}; x_{ij}'\beta + H_j,\ \hat{\sigma}_\varepsilon^2)\ N\left(H_j; \widehat{H}_j, \sigma_{H_j}^2\right) N(\beta; \hat{\beta}, v_\beta)\, dH_j\, d\beta$. Integrating out H_j gives

$\int N(Y_{ij}\,; x_{ij}{}'\beta\ + \widehat{H}_j\,, \widehat{\sigma}_{\varepsilon}^2 +\ \sigma_{\widehat{H}_j}^2)\ \mathrm{N}(\beta; \hat{\beta}, v_{\beta})d\beta$. Integrating out β then gives

$$[Y_{ij} \mid data,\, \widehat{\sigma}_H^2, \hat{\sigma}_{\varepsilon}^2] =\ \mathrm{N}(x_{ij}{}'\hat{\beta}\ + \widehat{H}_j\,, \hat{\sigma}_{\varepsilon}^2 +\ \sigma_{\widehat{H}_j}^2 +\ x_{ij}\, v_{\beta}\, x_{ij}^{'})$$

We therefore compute the probability of malnutrition as

$$P\big(Y_{ij}\ < -2\big|\hat{\beta}, \widehat{H}_j, \hat{\sigma}_{\varepsilon}\big) =\ \phi\left(\frac{-2 - x_{ij}^{'}\hat{\beta} -\ \widehat{H}_j}{\sqrt{\left(\hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\widehat{H}_j}^2 +\ x_{ij}\, v_{\widehat{\beta}}\, x_{ij}^{'}\right)}}\right), \qquad (4.3)$$

where ϕ denotes the standard Normal cumulative distribution function.

Recall that a value below -2 in the standardised Z-score Y_{ij} is the standard reference for declaring a child as malnourished.

The individual child and household-level variances σ^2_{ε} and σ^2_{H} respectively obtained from our multilevel model are used as a summary measure of the degree of similarity between WAZ, HAZ or WHZ in children within the same household. We presented the measure of variance explained by household as the household level variance (σ^2_{H}) expressed as a percentage of the total variance ($\sigma^2_{\varepsilon} + \sigma^2_{H}$) from the model. Thus,

$$\rho_H = \{\sigma_H^2/(\sigma_H^2 +\ \sigma_{\varepsilon}^2)\} \times 100, \qquad (4.4)$$

The above Equation (4.4) can be interpreted as the measure of the degree of correlation between nutritional outcomes of children within same household and the greater the value of these coefficients, the higher the level of the correlation.

The choice of multilevel regression model presented in this appendix in Equation (4.1) is appropriate because we have data on individual children under-five nested within households. Not recognising the hierarchical structure of the data could lead to underestimation of the standard errors for the regression coefficients, which in turn would lead to spurious statistical significance and incorrect inference.

4.6.3 Online Supporting Information Appendix S3: Supplementary material relating to the logistic regression analysis of the association between self-reported birth size (large/average versus small) and actual birthweight.

Table 4.5 The effect estimate (β) for the associations between size of child at birth (outcome variable) and actual birthweight from logistic regression model (n=830*).

	Parameter Estimate (β)	Standard Error	P-value
Intercept	-3.75	0.67	2.14×10^{-8}
Actual birthweight (kg)	1.96	0.24	2.32×10^{-16}

*In the logistic regression, we could not use all the 1,163 children with actual birthweight as reported in the study limitation due to incomplete data on size of child at birth as perceived by the mother.

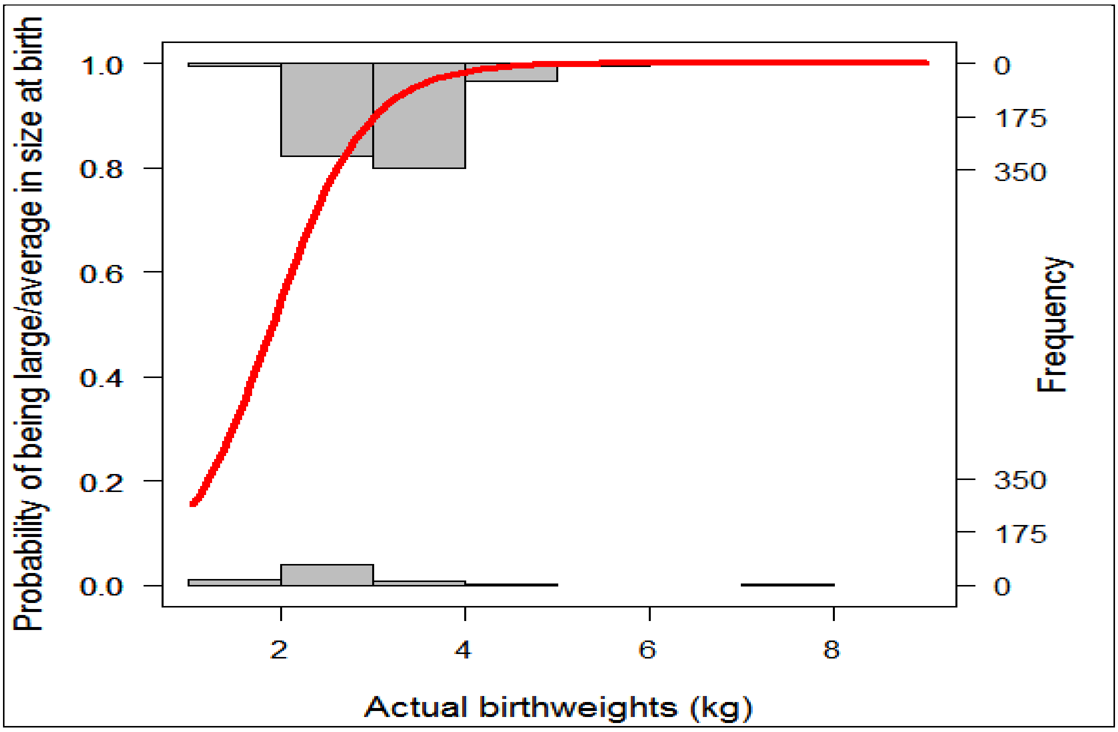


Figure 4.3 Probability plot of being born large/average versus small in size at birth against actual birthweight

4.7 Annex 1: STROBE checklist

Title and abstract

1a. Indicate the study's design with a commonly used term in the title or the abstract

This has been done in abstract section. The study was a cross sectional population based study.

1b. Provide in the abstract an informative and balanced summary of what was done and what was found

This has been done (page 53).

Introduction

2. Background/rationale. Explain the scientific background and rationale for the investigation being reported

This has been done. See page 55.

3. Objectives. State specific objectives, including any prespecified hypotheses

This has been done. See first and second paragraphs at page 56.

Methods

4. Study design. Present key elements of study design early in the paper

This has been done in the subsection "study population" (page 56)

5. Setting. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection

This has been done in the subsection "study population" (page 56)

6. Participants. Cross sectional study—Give the eligibility criteria, and the sources and methods of selection of participants

This has been done in the subsection "study population" (see pages 56-57)

7. Variables. Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable

This has been done in the subsections "outcome variables" (page 57) and "explanatory variables" (page 58, and page 73 in Table 4.4)

8. Data source/measurements. For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group

This has been done in the subsections "outcome variables" (page 57) and "explanatory variables" (page 58, and page 73 in Table 4.4)

9. Describe any efforts to address potential sources of bias

Source of possible bias have been described in the "study population" (page 56-57) detailing missingness in observations and have been stressed in study limitations (page 66).

10. Study size. Explain how the study size was arrived at

This has been done in the subsection "study population" (page 56-57).

11. Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why

This has been done in the subsection "statistical analysis" (page 58).

12. Statistical methods (a-e).

This has been done in the subsection "statistical analysis" (page 58-59).

Results

13. (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. (b) Give reasons for non-participation at each stage(c) Consider use of a flow diagram

This has been done in the section "results" (page 59) and in Tables (Tables 4.1-4.3). Flow diagram was not necessary because it was a one step cross sectional study.

14. Descriptive data. (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Cohort study—Summarise follow-up time (eg average and total amount)

This has been done in the subsection "sample characteristics" (page 59-61) and in Table 4.1. Follow-up time was not described because of the cross sectional nature of the study.

15. *Outcome data. Cross sectional study—Report numbers of outcome events or summary measures*

Summary measure of outcomes has been discussed at each specific point.

16. *Main results. (a) Report the numbers of individuals at each stage of the study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram*

This has been included in results and tables. Flow diagram was not necessary because it was a onetime cross sectional study.

17. *Other analyses. Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses*

Other analysis has been done with results presented in Table 4.5, page 76.

Discussion

18. *Key results. Summarise key results with reference to study objectives*

Key results have been summarised (see pages 64-66).

19. *Limitations. Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias*

Limitations have been discussed at pages 66 (last paragraph) and 68 (first paragraph).

20. *Interpretation. Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence*

A cautious interpretation considering other studies has been done (see pages 68-72).

21. *Generalisability. Discuss the generalisability (external validity) of the study results*

Generalisability has been discussed at each specific point.

Other information

22. *Funding. Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based*

This has been included at page 54.

4.8 Annex 2: Assessing severity of malnutrition by prevalence ranges

In this section we present further clarification on assessing severity of malnutrition by prevalence ranges as used in this Chapter.

The severity of malnutrition by prevalence ranges is based on the percentage of the population of children aged below 5 years old with Z-scores below -2 standard deviations (-2SD) cut-off point (de Onis and Blössner, 1997). For example, for stunting, underweight and wasting respectively, prevalence ranges of 20-29%, 10-19% and 5-9% is considered medium (moderate) form of severity for the three measures of malnutrition (de Onis and Blössner, 1997, Table 11, page 52).

References

Adekanmbi VT, Kayode GA, Uthman OA. (2013). Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis. *Maternal and Child Nutrition,* 9: 244-259.

Alom J, Quddus MA, Islam MA. (2012). Nutritional status of under-five children in Bangladesh: a multilevel analysis. *Journal Biosocial Science*, 44: 525-535.

Antwi S. (2008). Malnutrition: missed opportunities for diagnosis. *Ghana Medical Journal*, 4**2**:101-104.

Babatunde RO, Olagunju FI, Fakayode SB, Sola-Ojo FE. (2011). Prevalence and Determinants of Malnutrition among Under-five Children of Farming Households in Kwara State, Nigeria. *Journal of Agricultural Science,* 3: 173-181.

Black RE, Allen LH, Bhutta ZA, Caulfield LE, de Onis M, Ezzati M, et al. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. Lancet, 371: 243-260.

Blanchet NJ, Fink G, Osei-Akoto I. (2012). The effect of Ghana's National Health Insurance Scheme on health care utilisation. *Ghana Medical Journal,* 46:76-84.

Bomela NJ. (2009). Social, economic, health and environmental determinants of child nutritional status in three Central Asian Republics. *Public Health Nutrition,* 12:1871-1877.

de Onis M, Blössner M. (1997). *WHO Global Database on Child Growth and Malnutrition: Programme of Nutrition*. Geneva: WHO. [last accessed April 2013] (http:// whqlibdoc.who.int/hq/1997/who_nut_97.4.pdf).

de Onis M, Blössner M. (2003). The World Health Organization Global Database on Child Growth and Malnutrition: methodology and applications. *International Journal of Epidemiology,* 32: 518-526.

Food and Agriculture Organization (FAO). (2009). *Nutrition Country Profiles*. Rome: FAO, United Nations. [last accessed May 2015] (http://www.fao.org/ag/AGN/nutrition/gha_ en.stm).

Ghana Statistical Service, Ghana Health Service, ICF Macro. (2009). *Ghana Demographic and Health Survey 2008*. Accra: GSS, GHS, ICF Macro. [last accessed March 2013] (http://dhsprogram.com/pubs/pdf/FR221/ FR221%5B13Aug2012%5D.pdf).

Goldstein H. (2003). *Multilevel Statistical Models.* 3rd ed. London: Arnold.

Griffiths P, Madise N, Whitworth A, Matthews Z. (2004). A tale of two continents: a multilevel comparison of the determinants of child nutritional status from selected African and Indian regions. *Health and Place,* 10:183-199.

Hong R. (2006). Effect of multiple birth on infant mortality in Bangladesh. *Journal of Paediatrics and Child Health,* 42: 630-635.

Kabubo-Mariara J, Ndenge GK, Mwabu DK. (2009). Determinants of Children's Nutritional Status in Kenya: Evidence from Demographic and Health Surveys *Journal of African Economies,* 18:363–387.

Kandala NB, Madungu TP, Emina JB, Nzita KP, Cappuccio FP. (2011). Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter? *BMC Public Health,* 11:261.

Madise N, Mpoma M. (1997) Child malnutrition and feeding practices in Malawi. *Food and Nutrition Bulletin,* 18:190-201.

Mbago MC, Namfua PP. (1992). Some determinants of nutritional status of one- to four-year-old children in low income urban areas in Tanzania. *Journal of Tropical Pediatrics,* 38:299-306.

Melville B, Williams M, Francis V, Lawrence O, Collins L. (1988). *Determinants of Childhood Malnutrition in Jamaica*. Vol. 10. Tokyo: Food and Nutrition Bulletin, 43–48.

Mensah J, Oppong JR, Schmidt CM. (2010). Ghana's National Health Insurance Scheme in the context of the health MDGs: an empirical evaluation using propensity score matching. *Health Economics*, 19 Suppl:95-106.

Okoromah CA, Ekure EN, Lesi FE, Okunowo WO, Tijani BO, Okeiyi JC. (2011). Prevalence, profile and predictors of malnutrition in children with congenital heart defects: a case-control observational study. *Archives of Disease in Childhood*, 96: 354-360.

ORC Macro. (2005). Nutrition of Young Children and Mothers in Ghana**:** Findings from the 2003 Ghana Demographic and Health Survey. Calverton, Maryland, USA: Africa Nutrition Chartbooks.

R Core Team. (2014). *R: A language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. [last accessed June 2014] (https://www.r-project.org/).

Self SG, Liang K-Y.(1987). Asymptotic properties of maximum likelihood estimators and likelihood-ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82:605–610.

Steele F, Goldstein H, Rao C, Sinharay S. (2007). Multilevel Models in Psychometrics. *Psychometrics*, 26:401-420.

The World Bank. (2006). *Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action*. Washington DC: The International Bank for Reconstruction and Development/The World Bank. [last accessed August 2014] (http://siteresources.worldbank.org/ NUTRITION/Resources/281846-1131636806329/ NutritionStrategyOverview.pdf).

The World Bank. (2014). *World Development Indicators*. Washington, DC: The International Bank for Reconstruction and Development/The World Bank. [last accessedMay2014] (http://data.worldbank.org/sites/default/files/wdi-2014-book.pdf).

UNICEF, WHO.The World Bank. (2012). Joint Child Malnutrition Estimates: Levels and Trends. Geneva: WHO. [last accessed April 2014] (http://www.who.int/nutgrowthdb/estimates/ en/).

UNICEF-Ghana. (2008). Issues facing children in Ghana: UNICEF. [last accessed June 2014] (http://www.unicef.org/wcaro/Countries_1743.html).

UNICEF. (2006). The State of the World's Children. New York: UNICEF. [last accessed February 2014] (http:// www.unicef.org/sowc06/pdfs/sowc06_fullreport.pdf).

Uthman O. (2007). A multilevel analysis of the effect of household wealth inequality on under-five child under-nutrition: evidence from the 2003 Nigeria demographic and health survey. *The Internet Journal of Nutrition and Wellness*, 6:1–8.

Van de Poel E, Hosseinpoor AR, Speybroeck N, Van Ourti T, Vega J. (2008). Socioeconomic inequality in malnutrition in developing countries. *Bulletin of the World Health Organization*, 86: 282-291.

WHO. (2005). *Handbook: Integrated Management of Childhood Illness (IMCI)*. Geneva: WHO. [last accessed August 2015] (http://apps.who.int/iris/bitstream/10665/42939/1/9241546441.pdf).

WHO. (2010). *WHO Anthro for Personal Computers, Version 3.2.2, 2011: Software for Assessing Growth and Development of the World's Children*. Geneva: WHO. [last accessed April 2013] (http://www.who.int/childgrowth/software/en/).

WHO Multicentre Growth Reference Study Group. (2006). *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. Geneva: WHO. [last accessed April 2013] (http://www.who.int/childgrowth/standards/technical_report/en/).

Chapter 5 Paper 2: Modelling Spatial Variations in the Risk of Childhood Malnutrition among Under-Five Children in Ghana

This chapter is based on the following paper:

Aheto JMK, Taylor BM, Diggle PJ, Keegan TJ (2016). Modelling Spatial Variations in the Risk of Childhood Malnutrition among Under-Five Children in Ghana. Under review in *Statistics in Medicine*.

Abstract

Background: Childhood malnutrition varies geographically. While there are few studies investigating childhood malnutrition and its determinants in Ghana, spatial variation in under-five children's nutritional outcomes across the country remains unstudied. The present study seeks to fill this gap: we explore spatial variation in under-five child nutritional outcomes in Ghana while accounting for important risk factors; our goal is to help in the targeting of scarce resources and in the optimisation of public health policy intervention strategies in the country.

Methods: This study uses data from the nationwide Ghana Demographic and Health Survey (GDHS) conducted in 2008. Analyses were conducted on 2,051 under-five children residing in 393 geographic locations with eligible anthropometric measurements, using spatial mixed models. We produce predictive maps for the risk of childhood malnutrition over Ghana.

Results: Our results showed that healthy vegetation was associated with improved nutritional outcomes for children while high elevation and residing in rural areas are associated with increased risk of poor nutrition. Our predictive maps indicated that children from parts of Upper East, Eastern, Central, Western and Ashanti regions had the highest risk of malnutrition. Other relatively high risk areas include parts of Northern, Brong Ahafo, Volta and Upper West regions.

Conclusion: Under-five childhood malnutrition in Ghana is spatially structured. We suggest that in identified high risk areas there is an urgent need for public health intervention. This study has provided an improved understanding of spatial variation in childhood malnutrition situations across Ghana and its possible causes.

Keywords

Spatial modelling, Public health, Epidemiology, Biostatistics, Childhood malnutrition, Under-five children, Developing countries, Environmental epidemiology, Nutritional status, Malnutrition determinants

Funding

Funded by Faculty of Health and Medicine, Lancaster University, United Kingdom.

Except the funding of the study, the funder did not play any other role in this study including the submitted article.

5.1 Introduction

Improved nutrition is considered to be one of the foundations for better socio-economic development, especially in developing economies (The World Bank, 2006). Childhood malnutrition varies across geographic locations. Understanding the spatial pattern of its prevalence is important for the targeting of nutrition interventions amidst scarce public health resources in developing countries like Ghana.

The few studies into childhood nutrition in Ghana have mainly reported the overall prevalence of malnutrition as well as the effects of cultural, demographic and socioeconomic factors linked to under-five child malnutrition (Aheto et al., 2015; Amugsi et al., 2014; Antwi, 2008; Brugha and Kevany, 1994; Ghana Statistical Service et al., 2009) but spatial variation in under-five children's nutritional outcomes across the country remains unstudied.

Studies conducted in developing countries intended to examine spatial variation in childhood nutritional outcomes have established the importance of geographic location to improved understanding of childhood nutritional outcomes (Adekanmbi et al., 2013; Kandala et al., 2011; Khatab, 2010; Wand et al., 2012). Residential location can act as a marker for the personal, socioeconomic and environmental factors that influence people's health, nutritional status and access to health care services (Adekanmbi et al., 2013; Cressie, 1993; Diggle et al., 2002; Kandala et al., 2011). Spatial analysis therefore provides the needed tools to obtain an improved understanding of health outcomes by place and can contribute to the development of health interventions, treatments and prevention strategies (Nykiforuk and Flaman, 2011; Richardson et al., 2013).

In a previous study we conducted a non-spatial analysis, which examined risk factors associated with childhood malnutrition in Ghana using multilevel models (Aheto et al., 2015).

The main focus in the present paper is to apply spatial modelling techniques to explore spatial variation in childhood malnutrition measured on weight-for-age (WAZ), height-for-age (HAZ) and weight-for-height (WHZ) Z-scores over Ghana while adjusting for important environmental risk factors such as vegetation index, elevation and whether place of residence was rural or urban. Our goal is to produce predictive maps of the probability that a given child at any sampled or unsampled location in Ghana will be malnourished. These predictive maps can then be overlaid on the physical web map of Ghana to improve visualization of the prediction by relating them to the wider context.

5.2 Methods

5.2.1 Study population, design and sample

The present study uses data from the nationwide Ghana Demographic and Health Survey (GDHS) conducted in 2008 (Ghana Statistical Service et al., 2009).

Briefly, the GDHS employed a two-stage sampling design in which a total of 412 sample points (clusters) were selected nationwide in the first stage from a master sampling frame constructed from the 2000 Ghana Population and Housing Census. The second stage involved a selection of 30 households listed in each cluster (n=12,360). Data were not collected from one of the clusters. Sample weights were calculated taking into account household and individual non-responses in order to maintain representativeness of the sample. This resulted in 11,778 eligible households (only those households occupied at the time of the survey) being interviewed, with 5,175 and 6,603 located in urban and rural areas respectively. All women and men aged 15-49 and 15-59 years respectively in 50% of the households selected for the survey, and who were either usual residents or visitors present in the household on the night before the survey, were eligible for interview. Data were collected from representative samples of 4,916 women, from which data on 2,992 children under-five years old were collected. There were 756 children with incomplete information on the outcomes and exposures considered in this study. This study therefore focuses on 2,236 children. A detailed presentation of the GDHS study methods is available elsewhere (Ghana Statistical Service et al., 2009).

5.2.2 Outcome variable

The primary outcome of interest in our study was malnutrition among children under-five years of age. Malnutrition was assessed on three nutritional status indicators: height-for-age (HAZ), weight-for-age (WAZ) and weight-for-height (WHZ) Z-scores. The gender-specific Z-scores for each individual child for the three nutritional outcome measures were calculated using a macro provided by WHO (WHO, 2010). A child with a Z-score below minus two standard deviations for HAZ, WAZ or WHZ was declared as stunted, underweight or wasted, respectively (WHO Multicentre Growth Reference Study Group, 2006).

In accordance with published WHO methodology, we applied the WHO cut-off for out-of-range or biologically improbable values for Z-scores as outside $-6 \leq WAZ \leq 5$, $-6 \leq HAZ \leq 6$, and $-5 \leq WHZ \leq 5$, and removed these as outliers (WHO, 2010). This reduced our sample from 2,236 to 2,083 children. A further 32 children had no record of their geographical locations. We therefore analysed data on a total of 2,051 under-five children, whose geographical locations originated from 393 different clusters.

5.2.3 Explanatory variables

We adjusted for possible confounders of the relationship between location and nutritional outcome of children. Our selection of these possible confounders was informed by findings in the literature on the predictors of childhood nutrition in developing countries (Kandala et al., 2011; Kinyoki et al., 2015; Sobrino et al., 2014) and for which we had data for our study. These confounders were Normalized Difference Vegetation Index (vegetation index); elevation (km) and type of residence (urban=0 or rural=1). In this study, we focused on environmental, climatic and anthropogenic factors that might confound the relationship between location and nutritional outcome of children and for which we had data for our study. Presented in the Supporting Information Appendix S1 is the description of explanatory variables used in this analysis. We used the backward elimination method, which resulted in the retention of all three confounders in our final spatial models.

5.2.4 Statistical analysis

Our analysis explored whether children from different geographical locations exhibit different nutritional outcomes while simultaneously adjusting for potential confounders. The spatial model we used was

$Y_{ij} = x_{ij}^{'}\beta + S_j + \varepsilon_{ij}$ (4.1), with $S_j \sim N(0, \sigma_s^2\rho(d))$, and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$,

where Y_{ij} is the nutritional status (WAZ, HAZ or WHZ) for child i from community j, β is a vector of regression coefficients, x_{ij} is a vector of risk factors for child i in community j, S_j is the value of a correlated continuous spatial Gaussian process S at the spatial location of community j, σ_S^2 and ρ(d) are the marginal variance and the correlation matrix respectively associated with the spatial process S; since we are assuming S is stationary and isotropic, this correlation matrix is a function of the distance between communities only. The ε_{ij} are independent individual-level Gaussian residuals assumed to have variance σ_{ε}^{2}.

The correlation function $\rho(d)$ gives the correlation between locations distance d apart. Among competing correlation functions, a choice between Gaussian, Spherical and Exponential functions, the exponential function provided a good fit to the data. We therefore used the exponential correlation function, $\rho(d) = \exp(-d/ø)$, where d is the Euclidean distance between two locations and $ø > 0$ is the range or scale parameter controlling the rate at which correlation decreases with increasing distance, d. We fitted this model using the R (R Core Team, 2015) package 'spaMM' (Rousset and Jean-Baptiste, 2014). The standard likelihood ratio test was used to assess the statistical significance of the effects of the possible confounders. We computed the probability that a given child at an unsampled location will be malnourished based on our spatial model and produced predictive maps for these probabilities over Ghana (see Supporting Information Appendix S2). For detailed description of the statistical method used in this analysis, see Chapter 3 section 3.2.1 of this thesis.

5.3 Results

5.3.1 Sample characteristics

This study found that out of the 2,051 children considered in our analysis, 685 and 1,366 of them resided in urban and rural areas, respectively. Table 5.1 shows individual characteristics of the study population based on whether the individual is malnourished or not (eg. underweight or not underweight). For those who were underweight, 150 (55%) and 123 (45%) of them were males and females, respectively, while of those identified as stunted, 309 (53%) and 269 (47%) of them were males and females, respectively. Furthermore, for those classified as wasted, 83 (52%) and 78 (48%) of them were males and females, respectively.

Table 5.1 Percentage distribution of malnourished/not malnourished children by selected background characteristics (n=2051).

Characteristics	Underweight n (%)	Not underweight n (%)	Stunted n (%)	Not Stunted n (%)	Wasted n (%)	Not wasted n (%)
Sex						
Male	150 (55)	879 (49)	309 (53)	720 (49)	83 (52)	946 (50)
Female	123 (45)	899 (51)	269 (47)	753 (51)	78 (48)	944 (50)
Age (years)						
< 2	112 (41)	721 (41)	168 (29)	665 (45)	115 (71)	718 (38)
≥ 2	161 (59)	1,057 (59)	410 (71)	808 (55)	46 (29)	1,172 (62)
Breast-feeding duration						
≤ 6 months	16 (6)	209 (12)	15 (3)	210 (14)	33 (20)	192 (10)
> 6 months	257 (94)	1,569 (88)	563 (97)	1,263 (86)	128 (80)	1,698 (90)
Type of birth						
Single birth	255 (93)	1,728 (97)	555 (96)	1,428 (97)	154 (96)	1,829 (97)
Multiple birth	18 (7)	50 (3)	23 (4)	45 (3)	7 (4)	61 (3)
Diarrhoea episodes						
Had diarrhoea	76 (28)	374 (21)	144 (25)	306 (21)	52 (32)	398 (21)
No diarrhoea	197 (72)	1,404 (79)	434 (75)	1,167 (79)	109 (68)	1,492 (79)
Fever episodes						
Had fever	59 (22)	381 (21)	133 (23)	307 (21)	31 (19)	409 (22)
No fever	214 (78)	1,397 (79)	445 (77)	1,166 (79)	130 (81)	1,481 (78)
Size at birth						
Small	65 (24)	233 (13)	109 (19)	189 (13)	32 (20)	266 (14)
Large/average	208 (76)	1,545 (87)	469 (81)	1,284 (87)	129 (80)	1,624 (86)
Place of delivery						
Health care facility	128 (47)	972 (55)	281 (49)	819 (56)	74 (46)	1,026 (54)
None health care facility	145 (53)	806 (45)	297 (51)	654 (44)	87 (54)	864 (46)
Maternal education						
No formal education	117 (43)	663 (37)	233 (40)	547 (37)	70 (43)	710 (38)
Primary or higher	156 (57)	1,115 (63)	345 (60)	926 (63)	91 (57)	1,180 (62)
Household wealth						
Average/rich	91 (33)	835 (47)	201 (35)	725 (49)	58 (36)	868 (46)
Poor	182 (67)	943 (53)	377 (65)	748 (51)	103 (64)	1,022 (54)
Type of toilet facility						
Flush/pit	152 (56)	1,187 (67)	343 (59)	996 (68)	91 (57)	1,248 (66)
No toilet facility	121 (44)	591 (33)	235 (41)	477 (32)	70 (43)	642 (34)
Mother’s health insurance						
Has health insurance	93 (34)	743 (42)	199 (34)	637 (43)	59 (37)	777 (41)
No health insurance	180 (66)	1,035 (58)	379 (66)	836 (57)	102 (63)	1,113 (59)
Mothers’ BMI (kg/m^2)						
<18.5	32 (12)	121 (7)	50 (9)	103 (7)	21 (13)	132 (7)
≥18.5	241 (88)	1,657 (93)	528 (91)	1,370 (93)	140 (87)	1,758 (93)
Place of residence						
Urban	73 (27)	612 (34)	139 (24)	546 (37)	45 (28)	640 (34)
Rural	200 (73)	1,166 (66)	439 (76)	927 (63)	116 (72)	1,250 (66)

Out of those children identified as underweight, stunted and wasted, 112 (41%), 168 (29%) and 115 (71%), respectively, were aged below 2 years and for those classified as underweight, stunted and wasted, 161 (59%), 410 (71%) and 46 (29%), respectively, were aged 2 years or more. For those classified as underweight, stunted and wasted, 257 (94%), 563 (97%) and 128 (80%), respectively, were breastfed more than 6 months. Also, those children classified as underweight, stunted and wasted, 145 (53%), 297 (51%) and 87 (54%), respectively, were not delivered at health care facilities.

For those children identified as underweight, stunted and wasted, 117 (43%), 233 (40%) and 70 (43%), respectively, had mothers with no formal education. Furthermore, 182 (67%), 377 (65%) and 103 (64%) of the children who were underweight, stunted and wasted, respectively, belonged to poor households while 180 (66%), 379 (66%) and 102 (63%) of the children who were underweight, stunted and wasted, respectively, had mothers with no health insurance cover. For those children identified as underweight, stunted and wasted, 200 (73%), 439 (76%) and 116 (72%), respectively, were residing in rural areas.

5.3.2 Risks factors associated with WAZ, HAZ and WHZ for Spatial mixed models

In the results from the spatial analysis of WAZ, HAZ, and WHZ scores presented in Table 5.2, we found that vegetation index had no statistically significant association with HAZ , but had statistically significant, positive associations with WAZ and WHZ scores, suggesting that children living in communities where vegetation is richer have increased average Z-scores (i.e. improved average nutritional outcomes). Elevation had no statistically significant association with HAZ scores, but had statistically significant negative associations with WAZ and WHZ scores, suggesting that children living in communities at high elevations have lower average Z-scores (i.e. poorer nutritional outcomes). Place of residence status (rural/urban) had statistically significant association with WAZ, HAZ and WHZ scores. For all three outcomes, residing in rural areas was associated with poorer average nutritional outcomes compared to residing in urban areas.

Table 5.2 The effect estimate (β) for the associations between risk factors and child nutritional status (weight-for-age or WAZ, height-for-age or HAZ, and weight-for-height or WHZ) and spatial effect estimates for spatial mixed models (n=2051).

	Parameter estimates		
	WAZ (95% CI)	HAZ (95% CI)	WHZ (95% CI)
Intercept	-1.02 (-1.16, -0.88)	-1.13 (-1.32, -0.93)	-0.53 (-0.69, -0.38)
Vegetation (NDVI)	0.37 (0.17, 0.57)	-0.12 (-0.40, 0.16)	0.59 (0.36, 0.82)
Elevation (km)	-0.71(-1.14, -0.27)	-0.51 (-1.11, 0.09)	-0.56 (-1.05, -0.06)
Urban[a]			
Rural	-0.29 (-0.38, -0.19)	-0.33 (-0.46, -0.19)	-0.13 (-0.24, -0.01)
σ_s	0.32	0.84	0.36
σ_ε	0.99	1.14	1.11
ø (km)	8.18	7.52	9.73

[a]Reference category

5.3.3 Spatial variation in risk of malnutrition

Figure 5.1 shows the study locations used in the analysis. Figures 5.2, 5.3 and 5.4 show our predictive maps of the probability that a child at a given location in Ghana is malnourished.

Figure 5.1 Ghana map showing distribution of centroids (.) of the sampled clusters during the Ghana Demographic and Health Survey conducted in 2008

Figure 5.2 shows that the predicted malnutrition probability due to low WAZ ranges from about 0.04 to 0.28 across Ghana. The highest predicted probabilities, in the range 0.23 to 0.28, were observed in parts of Upper East region.

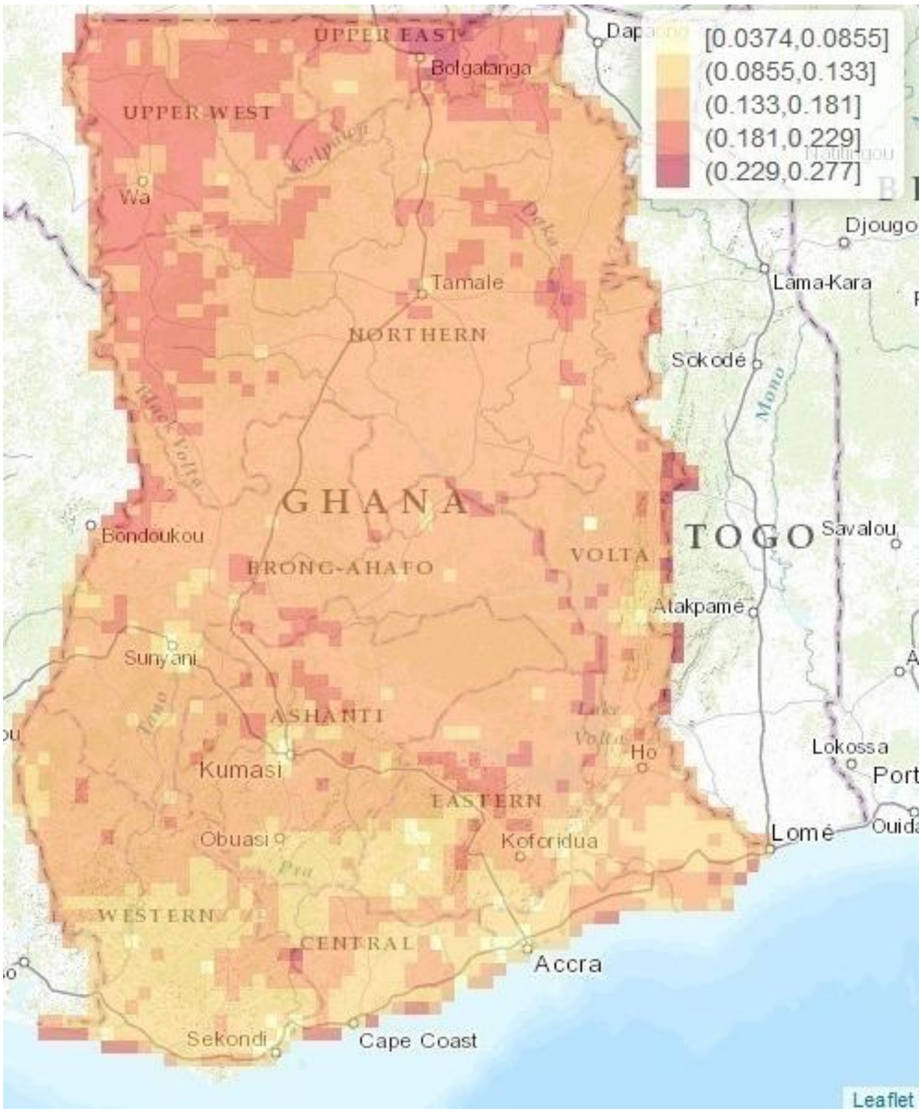


Figure 5.2 Spatial variations in probabilities of underweight in children <5 in Ghana

Figure 5.3 shows that the predicted malnutrition probability for HAZ ranges from about 0.07 to 0.68, with the highest predicted probabilities, ranging from 0.57 to 0.68, in parts of Upper East, Eastern, Central, Western and Ashanti regions. In other areas, including the Northern, Brong Ahafo, Upper West and Volta regions, predicted malnutrition probabilities range from about 0.32 to 0.56.

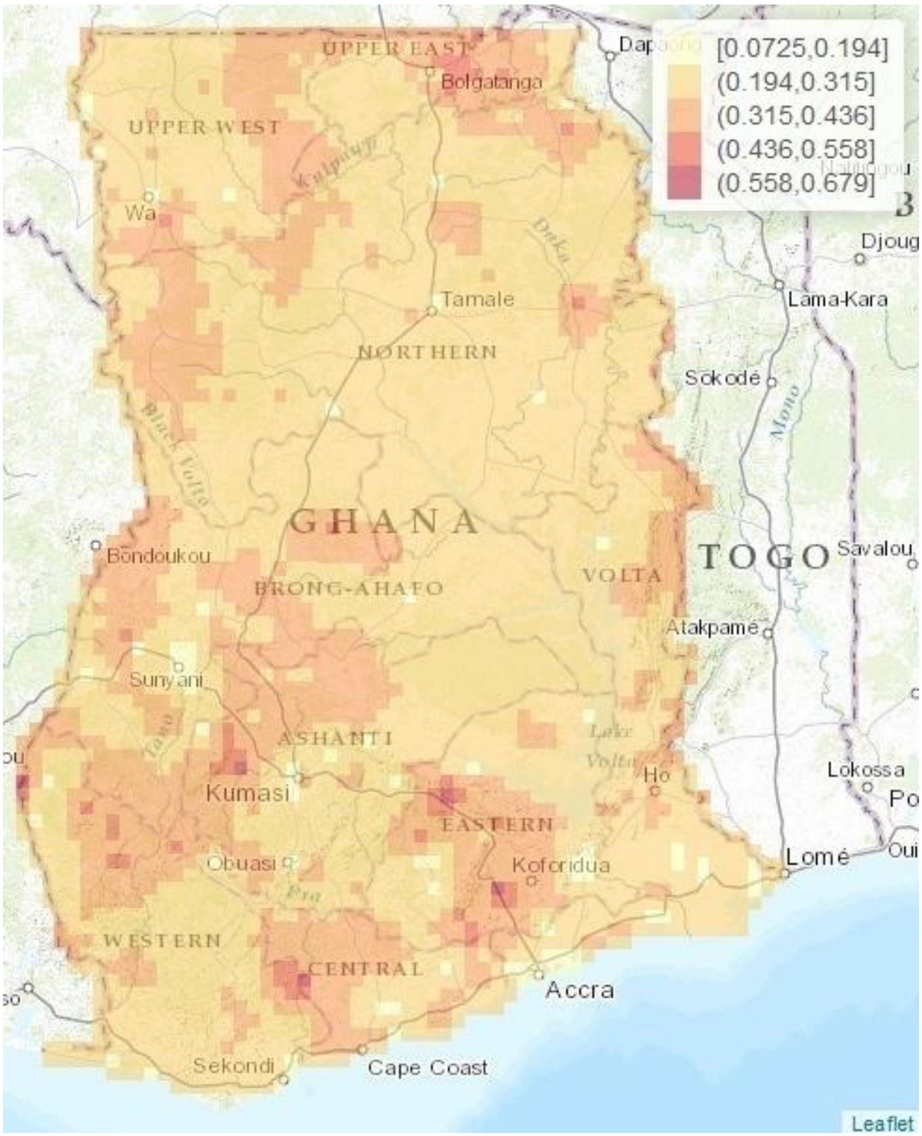


Figure 5.3 Spatial variations in probabilities of stunting in children <5 in Ghana

Figure 5.4 shows that the predicted malnutrition probability for WHZ scores ranges from about 0.02 to 0.15. The highest predicted malnutrition probabilities of about 0.13 to 0.15 were observed in parts of Upper West, Upper East and Northern regions.

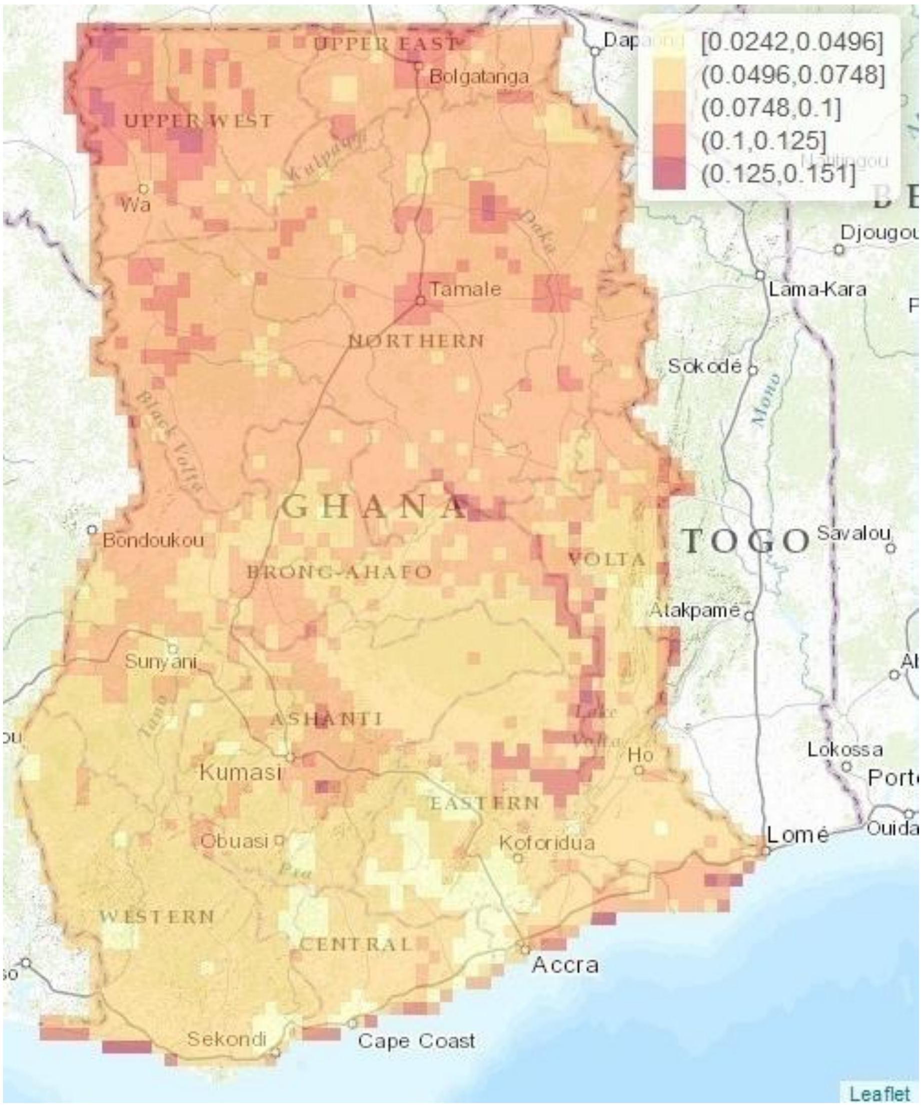


Figure 5.4 Spatial variations in probabilities of wasting in children <5 in Ghana

5.4 Discussion

Our main objective has been to investigate, model and illustrate the spatial variation in the prevalence of under-five child malnutrition in Ghana, as a contribution to better understanding of the problem of childhood malnutrition in the country.

We developed maps of the predicted probability of malnutrition; multiplication of these probabilities by 100 gives predicted prevalence (%). The methodology used to produce these maps takes into account the inherent uncertainties in model parameters and sampled data (see Supporting Information Appendix S2).

In our study the median predicted prevalence for underweight, stunting and wasting were 16%, 30% and 8% respectively. When we looked at spatial variation in the risk of malnutrition across Ghana, we observed that living in parts of Upper East, Eastern, Central, Western, Northern, Volta, Ashanti, Brong Ahafo and Upper West regions of the country was associated with relatively high probabilities of malnutrition. The Northern, Upper East and Upper West regions are the three regions of northern Ghana that are persistently the poorest among the ten regions in the country (Food and Agriculture Organization (FAO), 2009; Ghana Statistical Service, 2008; Van de Poel et al., 2007). These regions are predominantly agricultural and are dominated by smallholder farmers (five acres or less). The food crops produced are mostly for their own consumption, although some producers are also able to sell parts of their yield for income. Mostly, the farmers practice subsistence and rain-fed agriculture and farming activities in these regions. Accordingly, they are affected by recurrent floods and droughts, poor soil quality and increasingly erratic rainy seasons (Food and Agriculture Organization (FAO), 2009; World Food Programme (WFP), 2015). Thus, many people in these regions have varying levels of food available from year-to-year and a lack of variety in their diet, both of which could pose malnutrition challenges to children living in these regions (Food and Agriculture Organization (FAO), 2009). Another study conducted in Ghana among children under-five years of age also reported a higher prevalence of malnutrition in the three northern regions compared to other regions in the country (ORC Macro, 2005).

The Eastern, Central, Volta, Western, and Ashanti and Brong Ahafo regions are located at the southern part of Ghana and are less poor than the northern regions. Nevertheless, over 70% of the population in these regions are subsistence farmers and are affected by many of the same environmental issues as their counterparts in the north. Moreover, the risk of malnutrition in northern and southern parts of Ghana was higher in areas of high elevation. The range for elevation in Ghana is between 0 - 880 meters above the sea level and the risks were elevated in areas over about 200 meters.

From our results, environmental, climatic and anthropogenic factors as measured by vegetation index, elevation and rural/urban residence status appear markedly to affect the nutritional outcomes of children. We observed that vegetation index was positively associated with nutritional outcomes. This finding is consistent with previous studies, which have shown that children living in areas of healthy vegetation were at lower risk of malnutrition (Grace et al., 2012; Kinyoki et al., 2015). Elevation had a negative impact on children’s nutrition (Dang et al., 2004; de Meer et al., 1995). This finding supports previous studies, which found that children residing in areas of high elevation tended to exhibit increased risk of malnutrition and slow growth (Dang et al., 2004; de Meer et al., 1995). This could be due to physiological factors induced by high elevation observed to be related to low birthweight (Yip, 1987) which to some extent contributed to a reduction in postnatal growth and this finding appeared to provide an indirect support to our finding that high elevation increases the risk of child malnutrition. It is possible that genetic and socioeconomic background may confound the association we observed but we cannot rule out the possibility that elevation influences growth (de Meer et al., 1995).

Our finding support previous studies, which observed that residing in rural communities compared to urban communities increases the risk of childhood malnutrition (Fox and Heaton, 2012; Kandala et al., 2011; Smith et al., 2005). The likely reason for the difference is lack of access to modern health infrastructure and medical care services, coupled with lower socioeconomic status, and lack of, or inadequate knowledge on reproductive health and child care practices (e.g. feeding, child hygiene and safety) that, as in other developing countries, are more commonly observed in rural communities compared to urban communities (Food and Agriculture Organization (FAO), 2009; Fotso, 2007; Fox and Heaton, 2012; Ghana Statistical Service, 2008; Smith et al., 2005). The high malnutrition prevalence observed in both the northern and southern parts of Ghana could be related to poor access to health facilities and remoteness of most communities, especially in the rural areas that have poor transportation links.

The strengths of this study are the national coverage and representativeness of the study population, making the findings relevant to the wider population of Ghana. Our spatial model allows us to predict the prevalence of malnutrition in unsampled locations.

The limitations of this study include lack of access to some potentially relevant geographical or environmental-level covariates, such as distance to health facilities in the communities, which might help explain some of the observed spatial variation in the prevalence as well as incomplete data on some observations.

In the literature on malnutrition in sub-Saharan Africa, it is often the case that the prevalence of stunting is higher than for underweight, which in turn is higher than for wasting. We observed similar patterns in our study. One reason why the prevalence of wasting is usually lower than those for stunting and underweight is that wasting is sensitive to the impact on weight of infectious diseases, also to seasonal or temporal variations in food supply, and can therefore change rapidly in response to such factors (de Onis and Blössner, 1997).

Note that using our method of assessing malnutrition, even in a healthy, well-fed population of children aged <5 the prevalence would be shown as 2% as a direct consequence of the method of using z-score cut points. Sometimes, albeit rarely, this is subtracted from the estimated prevalence (de Onis and Blössner, 1997). Our results showed a predicted prevalence range of underweight, stunting and wasting of about 4% to 28%, 7% to 68% and 2% to 15%, respectively, over Ghana. Consequently the artefactual 2% prevalence would, if subtracted from the measured prevalence, have the greatest impact on the figures for wasting, which are the lowest of the three measures.

According to the World Health Organization (WHO), malnutrition in children is a result of multiple factors, often related to food quality, insufficient food intake, and severe and recurrent infectious diseases, or combinations of these factors. These conditions, in turn, are closely linked to the overall standard of living and whether a population can meet its basic needs, such as access to food, housing and health care. Thus, growth assessment not only serves as a means for evaluating the health and nutritional status of children but also provides an indirect measurement of the quality of life of an entire population (de Onis and Blössner, 1997).

This is the first study in Ghana to apply spatial models to investigate spatial variations in the risk of childhood malnutrition simultaneously controlling for important environmental, climatic and anthropogenic factors such as vegetation, elevation and rural/urban residence status.

The study identified communities where children are more at risk of malnutrition to inform efficient nutrition interventions in the country, the first of its kind in the country. Also, since public health resources are limited in Ghana, the spatial modelling approach adopted in this study will help optimize the limited resources available in the country; the limited resources will be spent on those communities that needed it most unlike the current nutrition interventions conducted on mass basis which are less or not supported with sound data and analysis, and could not reach everyone affected since the public health resources available in the country are limited and cannot support universal interventions.

Policy and intervention programmes targeting these higher-risk areas could be designed to include specific policies on childhood nutrition and health. There is a need to institute pragmatic and actionable strategies backed by a strong political will to develop and support the regions to become major food baskets and growth poles through adequate exploitation of the abundant rich natural resources available in these regions. Community-based nutrition surveillance systems in the form of nutrition evaluation and monitoring programs should be introduced and used to monitor the nutritional outcomes and health of children under-five years of age regularly, so as to provide regular up-to-date nutrition information on the under-five population to aid timely and relevant nutrition intervention strategies. Greater consideration should be given to the high-risk areas identified. Issues relating to environmental and climatic conditions should also contribute to the formulation of public health nutrition intervention strategies. Further research on factors not considered in this study could help to explain our observed patterns of spatial variation in the risk of childhood malnutrition.

5.5 Conclusion

Our predictive maps indicate malnutrition prevalence in unsampled areas of a country or region, here Ghana. The same methodology for constructing predictive maps can be applied in other countries and/or to other continuously measured health outcomes. These maps may be used to help optimize public health policy intervention strategies aimed at improving nutrition and health in low-income settings.

5.6 Appendix

5.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis.

Table 5.3 Description of explanatory variables used in this analysis

Community/environmental characteristics
Elevation (km): a measure of altitude (height above sea level) and is continuous
Residence: 1= rural and 0= urban
Vegetation index (NDVI): a measure of a vegetation cover and is continuous

5.6.2 Supporting Information Appendix S2: Supplementary material related to the spatial model and constructing predictive maps for probabilities of malnutrition

Let Y_{ij} be the nutritional status (WAZ, HAZ or WHZ) for child *i* from location *j*. We analysed the data using the model

$$Y_{ij} = x_{ij}^{'}\beta + S_j + \varepsilon_{ij} \qquad (5.1)$$

with $S_j \sim N(0,\sigma_S^2\rho(d))$, and $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$,

where β is a vector of regression coefficients, x_{ij} is a vector of risk factors, S_j is the value of a correlated continuous spatial Gaussian process S at the spatial location j, σ_S^2 and ρ(d) are the marginal variance and the correlation matrix respectively associated with the spatial process S; since we are assuming S is stationary and isotropic, this correlation matrix is a function of the distance between communities only. ε_{ij} are mutually independent individual-level residuals assumed to have zero-mean and variance σ_ε^2.

Among competing correlation functions, a choice between Gaussian, Spherical and Exponential functions, the exponential correlation function provided a good fit to the data. We therefore used the exponential correlation function, ρ(d) = exp(-d/ø), where d is the Euclidean distance between two locations and ø > 0 is the range or scale parameter controlling the rate at which correlation decreases with increasing distance, d.

We obtain estimates for $\hat{\beta},\ \hat{\sigma}_S^2\ and\ \hat{\sigma}_\varepsilon^2$ of the parameters in Equation (5.1) using maximum likelihood. For each location j, we estimated the minimum mean square error predictors $\hat{S}_j$ as the expectations of S_j conditional on our data and maximum likelihood parameter estimates.

Constructing predictive maps for probabilities of malnutrition

Our aim is to compute the probability that a given child at an unsampled location will be malnourished. To predict the nutritional status for a given child at an unsampled location, we have

$$\pi(Y_{ij}, x'_{ij} \mid data) = \int \pi(Y_{ij}, x'_{ij} \mid data)\ dx$$
$$= \int \pi(Y_{ij},\ x'_{ij} \mid data)\ \ \pi(x'_{ij} \mid data)\ dx$$ (5.2), and the expectation (E) of Y (nutritional outcome) given X (risk factors) is

$E_Y(Y) = E[E(Y|X)]$ (5.3) and its associated variance is

$V_Y(Y) = V[E(Y|X)]$ (5.4)

From our model, we required data on confounders used in our spatial model at the unsampled locations for which we don’t have these confounders in our data. We do have shapefiles containing the confounders (vegetation index and elevation) for the whole of Ghana which we can use for the unsampled locations but we did not have data on rural/urban residence for the unsampled locations. This means that data on rural/urban residence status is only available at the sampled locations for the sampled individuals but not available at the unsampled location for which we want to make our predictions. One possible solution will be to take a spatial average/mode of rural/urban values and use them to compute the probabilities. However, this approach has the following drawbacks:

1) It would not take into account the amount of individual variability in the data.

2) The effect of individual-level risk factors is not the same as the effects of group-level risk factors, which would lead to bias in resulting estimates (ecological fallacy).

To address this, we mapped the rural/urban covariate across Ghana and examined for any substantial spatial trend. We observed that the rural/urban covariate varied spatially by visual inspection, and is an important effect to account for in our predictions. To incorporate the rural/urban covariate in our predictions, we constructed a rural/urban covariate from the individual level data set. We used overlay operations to find which individuals lived in which of the prediction grid cells and determined a rural/urban covariate for each of these cells by majority voting; that is if there were more individuals in urban compared with rural settings in the prediction grid cell, we would select "urban" as the covariate value for that cell. This approach is similar to how elections are conducted to elect, say, members of parliament (MPs) to represent a particular constituency (here, predicted grid cell) in which the candidate receiving the majority of the valid votes cast will emerge as the winner and will be the one to represent the said constituency (here, predicted grid cell) in parliament. This process was repeated for all the grid cells covering the whole of Ghana. We used the newly created rural/urban covariate from our data set to provide predictions continuously over the whole of Ghana. The results in Equations (5.3) and (5.4) were used for the computation of probabilities of malnutrition.

Using ϕ to represent the standard normal cumulative distribution function, the probability that a given child will be malnourished at an unsampled location is therefore computed as

$$P\left(Y_{ij} < -2\right) = \phi\left(\frac{-2 - E_Y(Y)}{\sqrt{V_Y(Y)}}\right), \qquad (5.5)$$

where $E_Y(Y)$ is the predicted mean and $V_Y(Y)$ is its associated variance. Recall that a value below -2 in the standardised Z-score Y_{ij} is the standard reference for declaring a child as malnourished.

5.7 Annex 1: STROBE checklist

Title and abstract

1a. Indicate the study's design with a commonly used term in the title or the abstract

This has been done in abstract subsections. The study was a cross sectional population based study.

1b. Provide in the abstract an informative and balanced summary of what was done and what was found

This has been done (see page 84).

Introduction

2. Background/rationale. Explain the scientific background and rationale for the investigation being reported

This has been done. See pages 86-87.

3. Objectives. State specific objectives, including any prespecified hypotheses

This has been done. See second paragraph at page 87.

Methods

4. Study design. Present key elements of study design early in the paper

This has been done in the subsection "study population, design and sample" (page 87-88)

5. Setting. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection

This has been done in the subsection "study population, design and sample" (pages 87-88)

6. Participants. Cross sectional study—Give the eligibility criteria, and the sources and methods of selection of participants

This has been done in the subsection "study population, design and sample" (pages 87-88)

7. Variables. Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable

This has been done in the subsections "outcome variables" (page 88) and "explanatory variables" (page 89, and page 103 in Table 5.3)

8. Data source/measurements. For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group

This has been done in the subsections "outcome variables" (page 88) and "explanatory variables" (page 89, and page 103 in Table 5.3)

9. Describe any efforts to address potential sources of bias

Source of possible bias have been included in the "study population, design and sample" (pages 87-88) detailing missingness in observations and have been stressed in study limitations (page 100).

10. Study size. Explain how the study size was arrived at

This has been done in the subsection "study population, design and sample" (pages 87-88).

11. Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why

This has been done in the subsection "statistical analysis" (page 89).

12. Statistical methods (a-e).

This has been done in the subsection “statistical analysis” (pages 89-90).

Results

13. (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. (b) Give reasons for non-participation at each stage(c) Consider use of a flow diagram

This has been done in the section “results” (pages 90 and 92) and in Tables (Tables 5.1-5.2) with detailed description in subsection “outcome variable”, last paragraph on page 88. Flow diagram was not necessary because it was a one step cross sectional study.

14. Descriptive data. (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Cohort study—Summarise follow-up time (eg average and total amount)

This has been done in the subsection “sample characteristics” (pages 90-92) and in Table 5.1. Follow-up time was not described because of the cross sectional nature of the study.

15. *Outcome data. Cross sectional study—Report numbers of outcome events or summary measures*

Summary measure of outcomes has been discussed at each specific point.

16. *Main results. (a) Report the numbers of individuals at each stage of the study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram*

This has been included in results, tables and graphs. Flow diagram was not necessary because it was a onetime cross sectional study.

17. *Other analyses. Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses*

No other analysis has been done.

Discussion

18. *Key results. Summarise key results with reference to study objectives*

Key results have been summarised (see page 97-99).

19. *Limitations. Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias*

Limitations have been discussed at page 100.

20. *Interpretation. Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence*

A cautious interpretation considering other studies has been done (see pages 99-102).

21. *Generalisability. Discuss the generalisability (external validity) of the study results*

Generalisability has been discussed at each specific point.

Other information

22. *Funding. Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based*

This has been included at page 85.

References

Adekanmbi, V. T., O. A. Uthman, and O. M. Mudasiru (2013). Exploring variations in childhood stunting in Nigeria using league table, control chart and spatial analysis: BMC Public Health, v. 13, p. 361.

Aheto, J. M. K., T. J. Keegan, B. M. Taylor, and P. J. Diggle (2015). Childhood Malnutrition and Its Determinants among Under-Five Children in Ghana, Paediatric and Perinatal Epidemiology, v.**29**, p. 552-561.

Amugsi, D. A., M. B. Mittelmark, A. Lartey, D. J. Matanda, and H. B. Urke (2014). Influence of childcare practices on nutritional status of Ghanaian children: a regression analysis of the Ghana Demographic and Health Surveys: BMJ Open, v. 4, p. e005340.

Antwi, S. (2008). Malnutrition: missed opportunities for diagnosis: *Ghana Medical Journal*, v. 4**2**, p. 101-104.

Brugha, R., and J. Kevany (1994). Determinants of nutrition status among children in the eastern region of Ghana: Journal of Tropical Pediatrics, v. 40, p. 307-11.

Cressie, N. A. (1993). Statistics for Spatial Data: New York, John Wiley & Sons.

Dang, S., H. Yan, S. Yamamoto, X. Wang, and L. Zeng (2004). Poor nutritional status of younger Tibetan children living at high altitudes: European Journal of Clinical Nutrition, v. 58, p. 938-46.

de Meer, K., H. S. Heymans, and W. G. Zijlstra (1995). Physical adaptation of children to life at high altitude: European Journal of Pediatrics, v. 154, p. 263-72.

de Onis, M., and M. Blössner (1997). WHO Global Database on Child Growth and Malnutrition, Geneva, WHO.

Diggle, P., R. Moyeed, B. Rowlingson, and M. Thomson (2002). Childhood malaria in the Gambia: A case-study in model-based geostatistics, Journal of the Royal Statistical Society: Series C (Applied Statistics), Series C (Applied Statistics), p. 493-506.

Food and Agriculture Organization (FAO) (2009). Nutrition Country Profiles-Ghana. [last accessed May 2015] (http://www.fao.org/ag/AGN/nutrition/gha_en.stm), Rome, FAO, United Nations.

Fotso, J. C. (2007). Urban-rural differentials in child malnutrition: trends and socioeconomic correlates in sub-Saharan Africa: Health Place, v. 13, p. 205-23.

Fox, K., and T. B. Heaton (2012). Child nutritional status by rural/urban residence: a cross-national analysis: Journal of Rural Health, v. 28, p. 380-91.

Ghana Statistical Service (2008). Ghana Living Standards Survey: Report of the Fifth Round (GLSS 5), Ghana, GSS.

Ghana Statistical Service, Ghana Health Service, and ICF Macro (2009). Ghana Demographic and Health Survey 2008, Accra, Ghana., GSS, GHS, ICF Macro.

Grace, K., F. Davenport, C. Funk, and A. Lerner (2012). Child malnutrition and climate in Sub-Saharan Africa: An analysis of recent trends in Kenya, Applied Geography, p. 405-413.

Kandala, N. B., T. P. Madungu, J. B. Emina, K. P. Nzita, and F. P. Cappuccio (2011). Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter?: BMC Public Health, v. 11, p. 261.

Khatab, K. (2010). Childhood malnutrition in Egypt using geoadditive Gaussian and latent variable models: American Journal of Tropical Medicine and Hygiene, v. 82, p. 653-63.
Kinyoki, D. K., J. A. Berkley, G. M. Moloney, N. B. Kandala, and A. M. Noor (2015). Predictors of the risk of malnutrition among children under the age of 5 years in Somalia: Public Health Nutrition, p. 1-9.
Nykiforuk, C. I., and L. M. Flaman (2011). Geographic information systems (GIS) for Health Promotion and Public Health: a review: Health Promot Practice, v. 12, p. 63-73.
ORC Macro (2005). Nutrition of Young Children and Mothers in Ghana**:** Findings from the 2003 Ghana Demographic and Health Survey, Calverton, Maryland, USA, Africa Nutrition Chartbooks.
R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria.
Richardson, D. B., N. D. Volkow, M. P. Kwan, R. M. Kaplan, M. F. Goodchild, and R. T. Croyle (2013). Medicine. Spatial turn in health research: Science, v. 339, p. 1390-2.
Rousset, F., and F. Jean-Baptiste (2014). Testing environmental and genetic effects in the presence of spatial autocorrelation, Ecography p. 781-790.
Smith, L.C., M. T. Ruel, and A. Ndiaye (2005). Why is child malnutrition lower in urban than in rural areas? Evidence from 36 developing countries: World Development, v. 33, p. 1285-1305.
Sobrino, M., C. Gutiérrez, A. J. Cunha, M. Dávila, and J. Alarcón (2014). Child malnutrition in children under 5 years of age in Peru: trends and determinants: Revista Panamericana de Salud Publica, v. 35, p. 104-12.
The World Bank (2006). Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action, Washington DC, The International Bank for Reconstruction and Development/The World Bank. Assessed on 08/08/2014 at http://siteresources.worldbank.org/NUTRITION/Resources/281846-1131636806329/NutritionStrategyOverview.pdf.
Van de Poel, E., A. R. Hosseinpoor, C. Jehu-Appiah, J. Vega, and N. Speybroeck (2007). Malnutrition and the disproportional burden on the poor: the case of Ghana: International Journal for Equity in Health, v. 6, p. 21.
Wand, H., N. Lote, I. Semos, and P. Siba (2012). Investigating the spatial variations of high prevalences of severe malnutrition among children in Papua New Guinea: results from geoadditive models: BMC Research Notes, v. 5, p. 288.
WHO (2010). WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children., Geneva, WHO.
WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development, WHO.
World Food Programme (WFP) (2015). Ghana: Current issues and what the World Food Programme is doing. [last accessed September 2015] (https://www.wfp.org/countries/ghana), Rome, WFP.
Yip R. (1987). Altitude and birth weight: Journal of Pediatrics, v. 111, p. 869–876.

Chapter 6 Paper 3: Modelling and Forecasting Spatio-temporal Variations in the Risk of Chronic Malnutrition among Under-Five Children in Ghana

This chapter is based on the following paper:

Aheto JMK, Taylor BM, Keegan TJ, Diggle PJ (2016). Modelling and Forecasting Spatio-temporal Variations in Risk of Childhood Malnutrition among Under-Five Children in Ghana. Under review in *Spatial and Spatio-temporal Epidemiology*.

Abstract

Background: Children's nutritional outcomes vary over place and time. Spatio-temporal variation in under-five-year-old children malnutrition remains unstudied in most developing countries like Ghana. In this study, we explore and forecast the spatio-temporal patterns in childhood chronic malnutrition among these children and identify areas at high risk, to help improve the understanding of spatio-temporal variation and better to target pre-emptive nutrition interventions. We also investigate the effect of maternal education on childhood malnutrition.

Methods: We used the quinquennial data for the 1993, 1998, 2003 and 2008 Ghana Demographic and Health Surveys. In total, we analysed data on 10,036 children residing in 1,516 geographic locations at the time of the surveys. A spatio-temporal model was fitted to the data and was used to produce predictive maps of spatio-temporal variation in the probability of stunting over Ghana.

Results: The study found that out of the 10,036 children, 3,306 (33 %) of them were stunted and that there was substantial spatio-temporal variation in the prevalence of stunting. Our forecast map identified children from parts of Northern and Western regions to be at the highest risk of stunting with the prevalence ranging from about 54 % to 65 %. Higher levels of mother's education were associated with decreased risk of being stunted.

Conclusion: Our spatio-temporal model captured variations in childhood stunting over place and time. Our method facilitates and enriches modelling and forecasting of future stunting prevalence to identify areas at high risk. Improving maternal education could be given greater consideration within an overall strategy for addressing childhood malnutrition.

Keywords

Spatio-Temporal modelling, Forecasting, Public health, Epidemiology, Biostatistics, Childhood chronic malnutrition, Under-five children, Developing countries, Environmental epidemiology, Nutritional status, Malnutrition determinants

Funding

Funded by Faculty of Health and Medicine, Lancaster University, United Kingdom.

Except the funding of the study, the funder did not play any other role in this study including the submitted article.

6.1 Introduction

Childhood malnutrition is associated with poor mental and physical development, and with increased risk of morbidity and mortality (Black et al., 2008; de Onis and Blössner, 2003; Pelletier and Frongillo, 2003; UNICEF et al., 2012). It has been established that morbidity and mortality increase exponentially with deteriorating nutritional status (Pelletier DL et al., 1993). Childhood malnutrition also prevents children from reaching their full growth potential (de Onis and Blössner, 2003; UNICEF et al., 2012; World Bank, 2006). Childhood malnutrition is also associated with economic and health complications in adult life and constitutes a significant public health challenge globally, especially in developing countries (de Onis and Blössner, 2003; Ezzati et al., 2002; ORC Macro, 2005; The World Bank, 2006; UNICEF et al., 2012).

Spatial and temporal variations in the effects on growth of different levels of childhood nutrition can be attributable to socioeconomic, environmental and policy-related factors. Spatio-temporal analysis of under-fives malnutrition could therefore help to improve our understanding of the problem and to identify better and more targeted nutritional interventions, which is crucial to public health policy given the limited resources available in developing countries. Spatio-temporal variation in malnutrition amongst under-five-year-old children remains largely unstudied in most developing countries such as Ghana.

In a previous study using the 2008 Ghana Demographic and Health Survey (GDHS), we examined spatial variation in the risk of childhood malnutrition in Ghana, and identified higher risk communities in the country. Results from that study showed substantial spatial variation in the risk of childhood malnutrition in the country. The present study explores the spatio-temporal patterns in the risk of childhood malnutrition and provides a forecast for the risk of chronic malnutrition in Ghana.

Childhood malnutrition is the cause of about 40% of deaths among children under the age of 5 years in Ghana (Antwi, 2008; ORC Macro, 2005; UNICEF-Ghana, 2008) and remains a major public health challenge in the country; the percentages of children aged less than 5 years who were stunted were 33% in 1993, 31% in 1998, 35% in 2003 and 28% in 2008 (Ghana Statistical Service et al., 2009).

In this paper we investigate the spatio-temporal patterns in the risk of childhood chronic malnutrition measured on height-for-age Z-scores (HAZ, a measure of stunting) among children aged less than 5 years while taking into account the cross-sectional nature of each survey and adjusting for potential risk factors. We produce forecast maps for spatio-temporal variations in stunting over Ghana and identify areas at highest risk of stunting. We also evaluate the impact on the prevalence of malnutrition that would take place if we could change each child's covariates in certain ways. Specifically, we consider a policy change with regards to maternal education. In 2008, 55% of mothers had less than 6 years of education. To test the effect of increasing the level of maternal education we forecast what would have happened had each mother with less than 6 years education received exactly 6 years of education.

6.2 Methods

6.2.1 Study population and design

This study uses four data-sets from the quinquennial Ghana Demographic and Health Surveys (GDHS), conducted in 1993, 1998, 2003 and 2008 (Ghana Statistical Service et al., 2009; Ghana Statistical Service et al., 2004; Ghana Statistical Service (GSS) and Macro International Inc (MI), 1994, 1999). These data-sets included the same variables in each survey iteration, used the same survey methodology and were conducted by the same organization. The data were provided by the Ghana Statistical Service and DHS Program. They contain information on the health of women and their children under 5 years of age, their geographic location and anthropometric measurements (weight and height) at the time of the survey, breastfeeding and infant feeding practices, childhood illness and mortality, domestic violence, fertility, awareness and use of family planning. Anthropometric measurements of female respondents of reproductive age (15-49 years) and children under 5 years of age were collected from only those households selected for the individual interviews.

The GDHS used a two-stage cluster sampling design for each of the four surveys. The 1984 Ghana Population Census provided the master sampling frame from which the clusters or the sampling points were selected for the 1993 and 1998 GDHSs, whilst the 2000 Ghana Population and Housing Census (GPHC) provided the master sampling frame for the selection of the clusters for the 2003 and 2008 GDHSs. The clusters were selected with probability proportional to the number of households listed in each cluster. The selection of the clusters was followed by a complete listing of all households in the selected clusters, which provided the sampling frame for the selection of households in the second stage. Different samples and sample sizes were used in each of the four surveys, i.e. a serial cross-sectional rather than a longitudinal design. The data therefore have a hierarchical structure, with households nested within clusters and children nested within sampled households.

6.2.2 Geographical data

The 1993, 1998, 2003 and 2008 GDHSs identify the geographical locations of households with the centroid of the cluster from which they were selected. This resulted in a total of 319, 395, 402 and 400 distinct geographical coordinates for 1993, 1998, 2003 and 2008 GDHS data-sets, respectively. Data were not collected from one of the clusters sampled in the 2008 survey. Geographic coordinates were not recorded for 81, 10, 5 and 11 clusters in the 1993, 1998, 2003 and 2008 surveys, respectively, and these clusters were therefore excluded from our analysis. Out of the combined number of 1,623 clusters used across the surveys, 399 featured in all the 4 surveys, 1 featured in three out of four, 12 on two out of four. We used the 1,516 unique centroid cluster locations in our spatio-temporal analysis.

6.2.3 Outcome variable

The three most commonly used nutritional status measures for under-five children are weight-for-age (WAZ), weight-for-height (WHZ) and height-for-age (HAZ) Z-scores. We use the HAZ score, which is considered to be the most stable of the three as it is not easily influenced by temporal variations in food supply or impact of severe and recurrent diseases, and captures multiple dimensions of children's development, health and the environment in which they reside (de Onis and Blössner, 1997; Pradhan et al., 2003; WHO, 1995). Amongst these three measures, the HAZ score also gives the highest national prevalence of malnutrition in each of the survey years (Aheto et al., 2015; Ghana Statistical Service et al., 2009). A child is classified as stunted or chronically malnourished if their HAZ score is below -2 standard deviations from the median of the reference population (WHO Multicentre Growth Reference Study Group, 2006).

6.2.4 Explanatory variables

Months of breastfeeding, child's age, mother's educational attainment and household size have previously been identified in the literature to be predictive of under-five nutritional outcomes in developing countries (Adekanmbi et al., 2013a; Aheto et al., 2015; Alom et al., 2012; Babatunde et al., 2011; Bomela, 2009; Kabubo-Mariara et al., 2009; Kandala et al., 2011) and were available in the data from all four surveys. Presented in Supporting Information Appendix S1 is the description of explanatory variables used in this analysis. For our data, the backward elimination method selected these four covariates for inclusion in our final spatio-temporal model.

6.2.5 Numbers of eligible participants

In accordance with WHO guidelines, eligible HAZ scores are those in the range -6 to 6; values lying outside this range were removed as outliers (WHO, 2010). The numbers of children with incomplete data on height and age, ineligible HAZ scores or unidentifiable geographical coordinates in the 1993, 1998, 2003 and 2008 data-sets were 282, 562, 770 and 688, respectively. This resulted in 10,036 children who could be included in our analysis, consisting of 1,922, 2,736, 3,074 and 2,304 in 1993, 1998, 2003 and 2008, respectively (see Supporting Information Appendix S2 for details). Response rates from occupied households were 98.4%, 99.1%, 98.7% and 98.9% for 1993, 1998, 2003 and 2008, respectively (Ghana Statistical Service et al., 2009; Ghana Statistical Service et al., 2004; Ghana Statistical Service (GSS) and Macro International Inc (MI), 1994, 1999).

6.2.6 Statistical analysis

The analysis needs to accommodate: (i) the fact that some of the sampled locations change between successive surveys; (ii) our wish to be able to obtain spatially continuous maps of the probability of malnutrition. In this section we first introduce our statistical model for HAZ, accounting for the changes in sampled locations over the four surveys. We then present the details of our method for producing predictive maps.

6.2.6.1 Model and Methods

Our dynamic linear model for the HAZ outcome on a continuous scale, denoted by Y, is given by the state and observation equations, respectively (1) and (2) below.

$$\theta_t = A\theta_{t-1} + BW_t, \qquad W_t \sim \mathrm{MVN}(0, \Omega) \qquad (1)$$

$$Y_t = M_t\theta_t + V_{t,} \qquad V_t \sim \mathrm{N}(0,\ \sigma_v^2 I_{n_t}) \qquad (2)$$

where n_t is the number of children included at time *t* and I_{n_t} is an $n_t \times n_t$ identity matrix. Also,

$$\theta_t = \begin{bmatrix} S_t \\ \beta_t \end{bmatrix},$$

$$S_t = \left[S_{1,t}, S_{2,t}, . . ., S_{1516,t}\ \right]^T$$

$$\beta_t = \left[\beta_{1,t}, \beta_{2,t}, . . ., \beta_{5,t}\right]^T$$

$$A = diag[rep(\alpha, 1516), rep(1,5)]$$

$$B = diag\left[rep\left(\sqrt{(1-\alpha^2}), 1516\right), rep(0,5)\right]$$

$$\Omega = \left[\begin{array}{c|c} \Sigma & 0_{1516\mathrm{x}5} \\ \hline 0_{5\mathrm{x}1516} & 0_{5\mathrm{x}5} \end{array}\right]$$

$$M_t = [D_t \mid X_t]$$

In the above expressions: rep*(x,y)* is a vector of length *y* whose elements are all equal to *x*; diag[*v*] is a diagonal matrix with the elements of *v* on the diagonal; $0_{x\times y}$ is a matrix of zeros of dimension *x* by *y*; Σ is a matrix of dimension 1516 × 1516 whose *(i,j)* element is the covariance between the centroid of the *i*th cluster and the centroid of the *j*th cluster, which we model as $\sigma^2\exp(-d_{ij}/\phi)$, where d_{ij} is the distance between cluster *i* and cluster *j* (see below for justification); the (*i*,*j*) element of the n_t × 1516 matrix D_t is 1 if child *i* at time *t* was from cluster *j* and zero otherwise; the n_t × 5 matrix, X_t, is the design matrix for all observations at time t; $\alpha \in [0,1]$; and Y_t is a vector of observations at time t of length n_t.

This formulation is a linear state-space model (Kalman, 1960), whose state vector, θ_t, includes the spatio-temporal random effects, S_t, and survey-specific covariate effects, β_t. Note that Ω is not a valid covariance matrix since it has determinant zero. However, because our initial conditions for the state vector *are* valid, the Kalman filtering recursions can nevertheless be implemented with Ω defined as above. With α, A and B defined as above, the implication for the Spatiotemporal process S_t is that (unconditionally) if $S_{t-1} \sim N(0,\ \Sigma)$, then $S_t \sim N(0,\ \Sigma)$ also i.e S_t is stationary in time and has a separable covariance function.

The four model parameters are σ^2, σ_v^2, ϕ and α. Their interpretations are as follows: σ^2 is the unconditional variance of each $S_{i,t}$; σ_v^2 is the conditional variance of each element of Y_t given the state vector θ_t; ϕ controls the rate at which the correlation between the $S_{i,t}$ at different locations decays with increasing distance between them; α is the autocorrelation between $S_{i,t}$ and $S_{i,t+1}$. We could have included the term $X_t\beta$ into the observation equation (2), but this would have considerably increased the computation cost.

To complete the model, we set the initial conditions of the filter to be Gaussian with mean, μ_0 and covariance Σ_0. This choice guarantees that the *time t* filtering distribution $[\theta_t|Y_{1:t}]$, where $Y_{1:t}$ denotes the set of all Y_s for *s = 1, 2, .., t*, is also Gaussian for all *t*. We chose the initial conditions as follows. For μ_0, we fitted a linear mixed effect model to the 1993 data-set using geographic location as a single, spatially and temporarily uncorrelated grouping variable, and set μ_0 as its maximum likelihood estimate under this working model. For Σ_0, we conducted a sensitivity analysis using the choices $\Sigma_0 =$

diag[rep(u, 1521)] with $u \in \{1, 10, 20, 50, 60\}$. All choices led to substantially the same results; we used u=10 in our main analysis.

We estimated the four parameters in our model, α, ϕ, σ^2, σ_v^2, using maximum likelihood. For the parameters ϕ, σ^2 and σ_v^2, we initialised the maximum likelihood procedure by conducting a variogram analysis of the 1993 data-set. We set the initial value of the autocorrelation parameter, α, to be 0.5, the mid-point of its support. The variogram analysis was also used to choose the functional form for ρ, a choice between the Exponential, Gaussian and Spherical correlation functions. We obtained 95% confidence intervals for our fixed effects from the filtering distribution; Table 6.1 gives β_t for the last time point, conditional on having observed data from all surveys up to and including 2008.

We then used the fitted model to compute the probability that a child at an unsampled location will be chronically malnourished (stunted) i.e for which HAZ < -2, and constructed predictive maps for these probabilities across Ghana. We also constructed forecast maps of the effect of varying levels of maternal education on the probabilities of chronic malnutrition across Ghana.

All computations were performed using the R package `miscFuncs', which implements the Kalman Filter and parameter estimation for Gaussian dynamic linear models (R Core Team, 2015; Taylor, 2015). For detailed review of the statistical method used in this analysis, see Chapter 3 sections 3.2.2 through 3.2.4 of this thesis.

6.2.6.2 Predicting the Risk of Stunting Across the Whole of Ghana

One of our aims is to forecast the probability that in the year of a future survey, here 2013, a child living at an arbitrary location will be stunted. In order to produce maps, we forecast the probability of stunting on a grid of points covering Ghana. This requires us to know or assume the values of all relevant covariates. For forecasting, we propose to draw from the empirical distribution of covariates at the most recent survey time point (i.e. 2008) and use these samples to construct forecast probabilities of stunting across Ghana. We make two assumptions. The first is that the distribution of the characteristics of the population in 2008, namely duration of breastfeeding, mothers' education level, household size and the children's age as measured in the survey will not change substantially in the time we are intending to forecast. The second is that the distribution of each of these population characteristics is approximately uniform over Ghana.

From our model, we can compute the one-step ahead forecast density of $[\theta_{t+1} | Y_{1:t}]$ exactly. Using the mean and variance of this density, we can then predict the distribution of the spatial random effect on the finely spaced grid covering Ghana. This is a multivariate Gaussian distribution with mean $S_G^{(t+1)}$ and variance $V_G^{(t+1)}$. To predict the probability of malnutrition on this grid, we again need to know or assume the values of individual-level covariates at each point on the grid. One possible solution would be to take a spatial average of risk factor values over a suitable neighbourhood of each grid point, and use these to compute the required probabilities. However, this approach has the following drawbacks:

1. It would not take into account the amount of individual variability in the data;
2. The effect of individual-level covariates is not the same as the effects of group-level covariates; ignoring this would lead to ecological bias in the resulting estimates.

We therefore propose to use individual level covariate data to capture the individual-level variability, but with a homogeneous distribution over space.

To obtain the forecast mean and variance of HAZ for 2013 on the grid, we used the covariates of the *i*th child in 2008, $X_{t+1}^{(i)}$, to compute the forecast mean and variance for a child with the same covariates over Ghana,

$$M^{(i)} = X_{i,t+1}\beta_{t+1} + S_G^{(t+1)}$$

$$V^{(i)} = (X_{i,t+1})^T V_\beta^{(t+1)} X_{i,t+1} + V_s^{(t+1)} + \hat{\sigma}_v^2,$$

where $V_\beta^{(t+1)}$ and $V_s^{(t+1)}$are the appropriate sub-matrices of $V_G^{(t+1)}$and $\hat{\sigma}_v^2$ is estimated nugget effect. Finally, we use equations (3) and (4) to combine these estimates and obtain the mean and variance of the forecast distribution:

$$\mathbb{E}_Y(Y_{t+1}|data) = \mathbb{E}_{X_{t+1}}[\mathbb{E}_{Y_{t+1}|X_{t+1}}(Y_{t+1}|X_{t+1}, data)], \qquad (3)$$

$$\mathbb{V}_Y(Y_{t+1}|data) = \mathbb{V}_{X_{t+1}}[\mathbb{E}_{Y_{t+1}|X_{t+1}}(Y_{t+1}|X_{t+1}, data)] + \mathbb{E}_{X_{t+1}}[\mathbb{V}_{Y_{t+1}|X_{t+1}}(Y_{t+1}|X_{t+1}, data)]. \qquad (4)$$

We then evaluate exceedance probabilities, i.e., *P(HAZ< -2)*, using the standard normal cumulative distribution function. In (3) and (4) we use Monte Carlo integration to evaluate the outer integral, selecting covariate values, $X_{i,t+1}$, by sampling from their empirical distribution in 2008.

6.2.6.3 Exploring the Effect of Policy Changes

A second aim is to produce predictive maps of the effect of varying levels of maternal education on the risk of childhood chronic malnutrition. We achieved this by modifying $M^{(i)}$ and $V^{(i)}$ above, potentially changing each child's covariates depending on whether that child's mother received (i) less than 6 years or (ii) less than 12 years education. In case (i), we changed all maternal education covariates to equal 6 years if the child's mother received less than 6 years of education. In case (ii), we changed all maternal education covariates to equal 12 years if the child's mother received less than 12 years of education. We again used equations (3) and (4) to obtain the mean and variance of the policy-modified predictive distribution.

6.3 Results

Out of the 10,036 children with eligible HAZ scores and geographical coordinates 3,306 (33%) were identified as being stunted. Table 6.1 presents the estimates of the fixed effects in our spatio-temporal model for HAZ score. Months of breastfeeding, child's age, and mother's years of education were statistically significant risk factors for HAZ score, whereas number of household members had no statistically significant association with HAZ score. Longer breastfeeding duration and older ages of children were negatively associated with HAZ score for children, while higher levels of mother's education were positively associated with HAZ scores for children.

Table 6.1 The effect estimate (β) for the associations between risk factors and child nutritional status (height-for-age Z-score or HAZ score) for spatiotemporal model (n=10,036)

	Parameter estimates for HAZ (95% CI)
Intercept	-1.288 (-1.421, -1.156)
Months of breastfeeding	-0.445 (-0.485, -0.404)
Child's age	-0.195 (-0.236, -0.154)
Mother's education	0.105 (0.070, 0.140)
Household size	0.008 (-0.024, 0.040)

Table 6.2 presents the parameters of the fitted spatio-temporal model for HAZ. The range, which measures the rate at which spatial correlation decays to zero with increasing distance, is about 95km. The estimated value of the autocorrelation between the successive values $S_{i,t}$ at the same location, α = 0.07. To formally test the effect of

the dependency between states, we used a likelihood ratio (deviance) test of the hypothesis that α=0. The deviance statistic was 3.6 on 1 degree of freedom, which is not conventionally significant (*p-value = 0.0578*). Nevertheless, we included the estimate $\hat{\alpha}$ in our spatio-temporal model to enable prediction of future stunting patterns.

Table 6.2 Parameters of the fitted spatiotemporal model for height-for-age Z-score (or HAZ)

Partial sill	$(\sigma^2)^A$	Nugget $(\sigma_v^2)^B$	Range $(\phi)^C$	α^D
Estimates	0.37	1.51	94.85 km	0.07

A= Is the unconditional variance of each $S_{i,t}$.
B= Is the conditional variance of each element of Y_t given the state vector θ_t.
C= Controls the rate at which the correlation between the $S_{i,t}$ at different locations decays with increasing distance between them.
D= Is the autocorrelation between $S_{i,t}$ and $S_{i,t+1}$.

Figure 6.1 presents the maps of the geographical variation in the probability that a child will be stunted in 1993, 1998, 2003 and 2008, using the same scales for all four years. The ranges of the predicted probabilities of stunting in 1993, 1998, 2003 and 2008 are about 0.23 to 0.52, 0.19 to 0.44, 0.37 to 0.65 and 0.17 to 0.37, respectively. The highest stunting probability range of about 0.56 to 0.66 was observed in parts of Northern and Western regions in the year 2003.

At the time of writing, the 2008 Ghana Demographic and Health Survey is the most recent survey data-set available to the public. To illustrate the performance of our fitted spatio-temporal model, we therefore used it to forecast the probabilities of stunting for the year 2013.

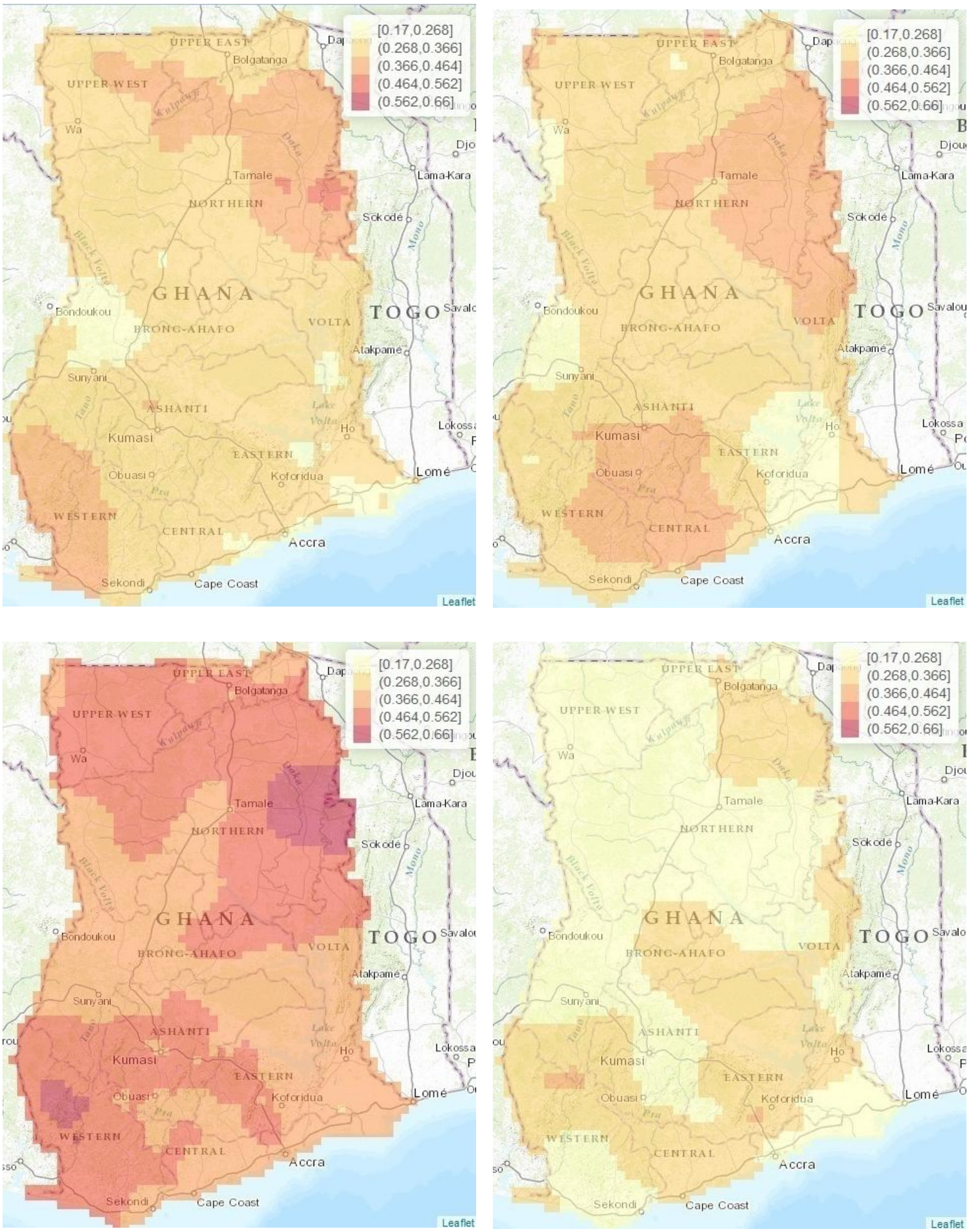


Figure 6.1 Probabilities of stunting in Ghana in 1993 (top left), 1998 (top right), 2003 (bottom left) and 2008 (bottom right) in children <5 years

Figure 6.2 shows our forecast map, using the same range as in Figure 6.1. The predicted stunting probability ranges between 0.36 and 0.65. Stunting probabilities in the range 0.54 to 0.65 were predicted in parts of Northern and Western regions of the country, while probabilities in the range 0.42 to 0.53 was observed in parts of Upper West, Upper East, Ashanti, Volta, Eastern, Brong Ahafo, Central and Greater Accra regions of the country. The geographical distribution in the risk of childhood chronic malnutrition observed in our forecast maps over Ghana is similar to those observed in each of the actual surveys.

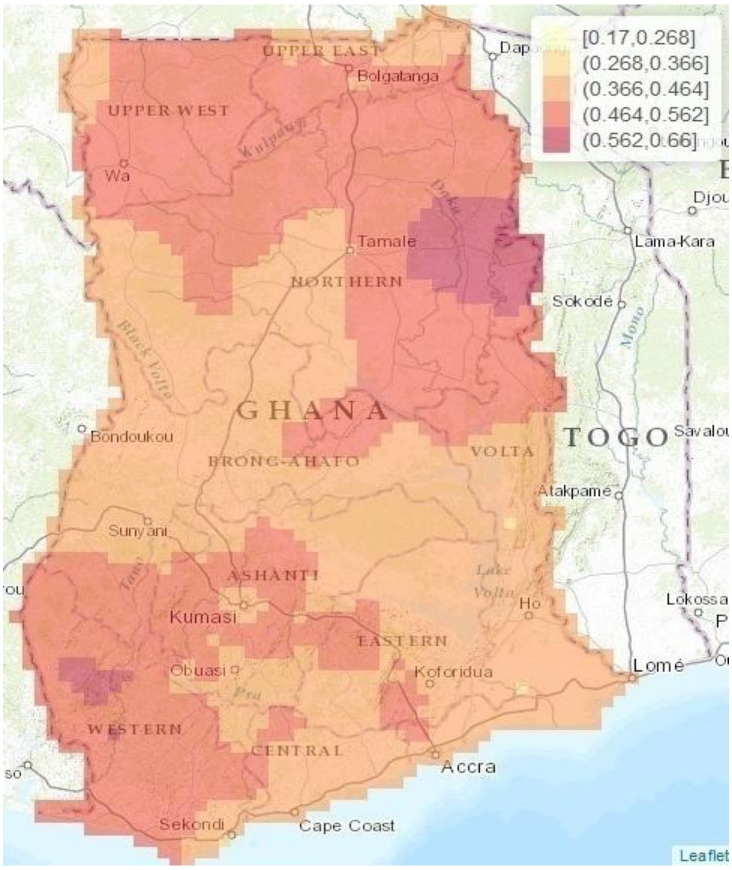


Figure 6.2 Overall probabilities of stunting forecast in Ghana in 2013 in children <5 years

We also investigated the effect of mother's education on the risk of childhood chronic malnutrition across Ghana by making predictions of childhood malnutrition under different levels of maternal education in the country. Figure 6.3 presents our predicted maps of probabilities of childhood stunting based on mothers who had a minimum of 6 years and 12 years of education; note that both panels use the same range but that this is different from the range used in Figures 6.1 and 6.2. The predicted stunting probability ranges from about 0.17 to 0.26.

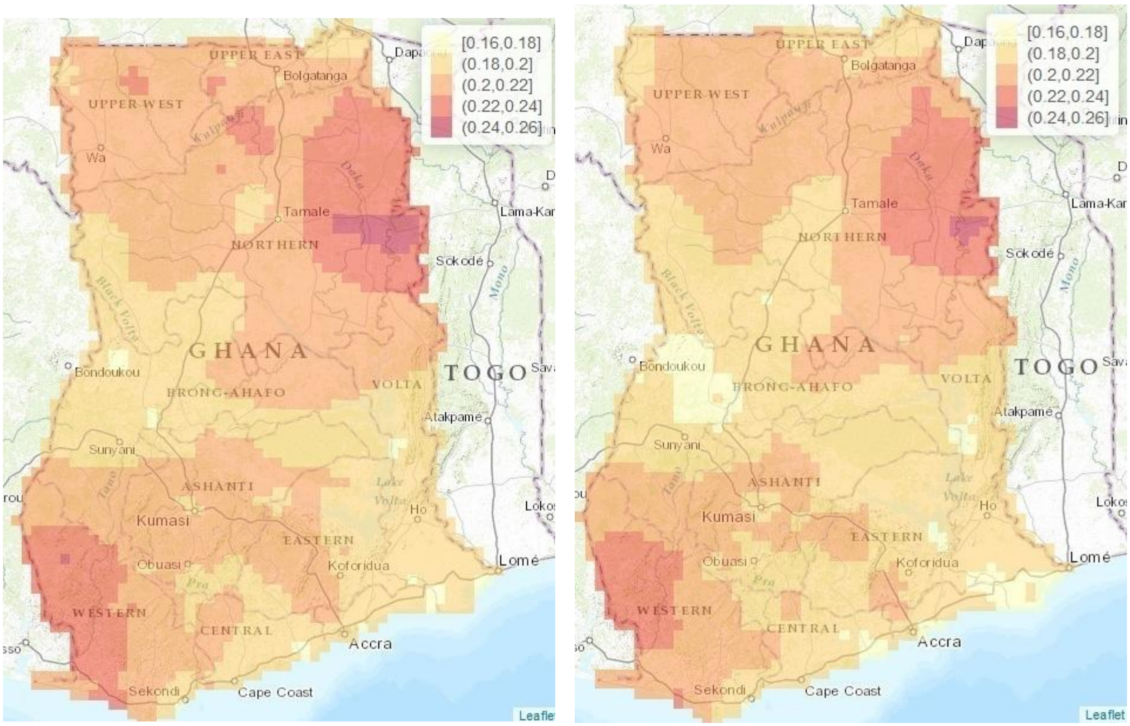


Figure 6.3 Effect of a minimum of 6 (left) and 12 (right) years of mothers' education on probabilities of stunting forecast in Ghana in 2013 in children <5 years

6.4 Discussion

In this study, our aim has been to contribute to a better understanding of childhood malnutrition problems in Ghana and to inform pre-emptive nutrition interventions. We constructed maps showing predicted probabilities of chronic malnutrition in children for the years 1993, 1998, 2003, 2008 and 2013. We have taken into account the uncertainties in our model parameters and in childhood characteristics implicit in these maps (see the methods section).

We found substantial spatial and temporal variations in the probabilities of chronic malnutrition in Ghana, with median predicted prevalence of 33%, 33%, 47% and 26% for the years 1993, 1998, 2003 and 2008, respectively. We observed that living in parts of Northern or Western regions was associated with relatively high probabilities of chronic malnutrition.

Our forecast median stunting prevalence for the year 2013 in the country was 46%; living in parts of Northern or Western regions was again associated with significantly and materially higher probabilities of chronic malnutrition. Among all the ten regions of Ghana, Northern region was among the three regions in the northern part of the country that are consistently the poorest; the other two are Upper East and Upper West (Food and Agriculture Organization (FAO), 2009; Ghana Statistical Service, 2008; Van de Poel et al., 2007). This could partly explain the high chronic malnutrition prevalence patterns observed in Northern region.

According to the ‘Feed the Future’ program, a United States Government global hunger and food security initiative, chronic malnutrition in the northern regions of Ghana is related to household poverty levels, inadequate sanitation facilities and poor infant-feeding practices, and leads to higher disease burden (Feed the Future, 2015). Our finding also supports a previous study conducted in Ghana among children aged less than 5 years, which reported higher prevalence of malnutrition in the northern part compared to the southern part of the country (ORC Macro, 2005; Van de Poel et al., 2007).

Children from mothers with higher levels of education were at less risk of malnutrition. This might be because a mother’s education is associated with knowledge of good practices for herself and her child (e.g. feeding and health seeking behaviour) but is also associated with household socioeconomic status and access to food (Kandala et al., 2011; ORC Macro, 2005; Van de Poel et al., 2007). In our predictive maps shown in Figure 6.3, increasing the level of maternal education was shown to reduce the prevalence of malnutrition throughout Ghana. However, policies and interventions aimed at improving the level of maternal education to combat childhood malnutrition must be supported with efforts to improve the general standard of living among households in the communities.

The parameter estimates from our spatio-temporal model characterise the nature of the spatio-temporal variations in nutritional outcomes of children. We observed a weak, and statistically non-significant, time-dependence between successive surveys, suggesting that a spatial model for each of the survey years separately could fit the data as well as our spatio-temporal model. However, one of our main aims in this paper is to forecast the risk of childhood chronic malnutrition, which is only possible through a spatio-temporal model. A longitudinal survey design, with repeated measurements in the same household, would undoubtedly have led to stronger correlation between successive surveys.

The study's strengths include representativeness and national coverage, which makes the findings relevant to the wider population of Ghana. In addition, our spatio-temporal model permitted us to borrow strength from similar previous surveys conducted in 1993, 1998, 2003 and 2008 to provide a forecast for stunting prevalence in the country in 2013.

The study's limitations include lack of access to other possible risk-factors such as data on interventions (nutrition, health care and socioeconomic impacts) and distance to health care facilities in the communities as well as missingness in observations.

Previous studies in developing countries like Ghana have shown that socioeconomic, cultural and demographic factors, feeding and care practices are associated with nutritional outcomes of children under five years of age (Adekanmbi et al., 2013a; Adekanmbi et al., 2013b; Aheto et al., 2015; Amugsi et al., 2014; Antwi, 2008; Bomela, 2009; Brugha and Kevany, 1994; Kandala et al., 2011). However, these studies did not explore or forecast spatio-temporal variations in nutritional outcomes.

Generally, lack of food availability, poverty, lack of dietary diversity, inadequate care and feeding practices coupled with inadequate or complete lack of access to health care services are among the major factors responsible for malnutrition in children (de Onis and Blössner, 1997; Food and Agriculture Organization (FAO), 2009; Van de Poel et al., 2007).

Using the same confounders, the risk of stunting is greatest in 2003 than other years. This may be due to random variations which could not be controlled for in this study. However, we did not rule out the possibility that the observed pattern in the risk of stunting in 2003 could be real and further research in this area could explore this issue.

Our results can be used to target public health nutrition interventions for high-risk communities so as to improve overall health and nutrition of Ghanaian children in two different ways. Firstly, as with our example of maternal education, our model enables prediction of the effectiveness of a public health intervention that changes the distribution of an included explanatory variable. Secondly, our maps can identify priority areas for implementation of an intervention that is known to be effective but is too expensive to implement universally.

6.5 Conclusion

Our forecast maps of probabilities of stunting for the year 2013 illustrate how our spatio-temporal modelling approach can be used as a tool for identifying high risk communities. This can help to prioritise and target nutrition and health policies that can promote effective and sustainable public health interventions in Ghana and other developing countries, with the overall aim of improving childhood nutrition and health.

We also intend to apply our current spatio-temporal modelling techniques to update our model and provide forecasts for the year 2019, as soon as the 2014 GDHS data-set is made available.

6.6 Appendix

6.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis.

Table 6.3 Description of explanatory variables used in the analysis

Child characteristics
Age: Age in years and is continuous
Breastfeeding duration: breastfeeding duration in months and continuous
Maternal/household characteristics
Mother's education: mother's education in years and is continuous
Household size: number of persons in a household and is continuous

6.6.2 Supporting Information Appendix S2: Supplementary material related to numbers of children included/excluded for each of the four data collection rounds used in the study.

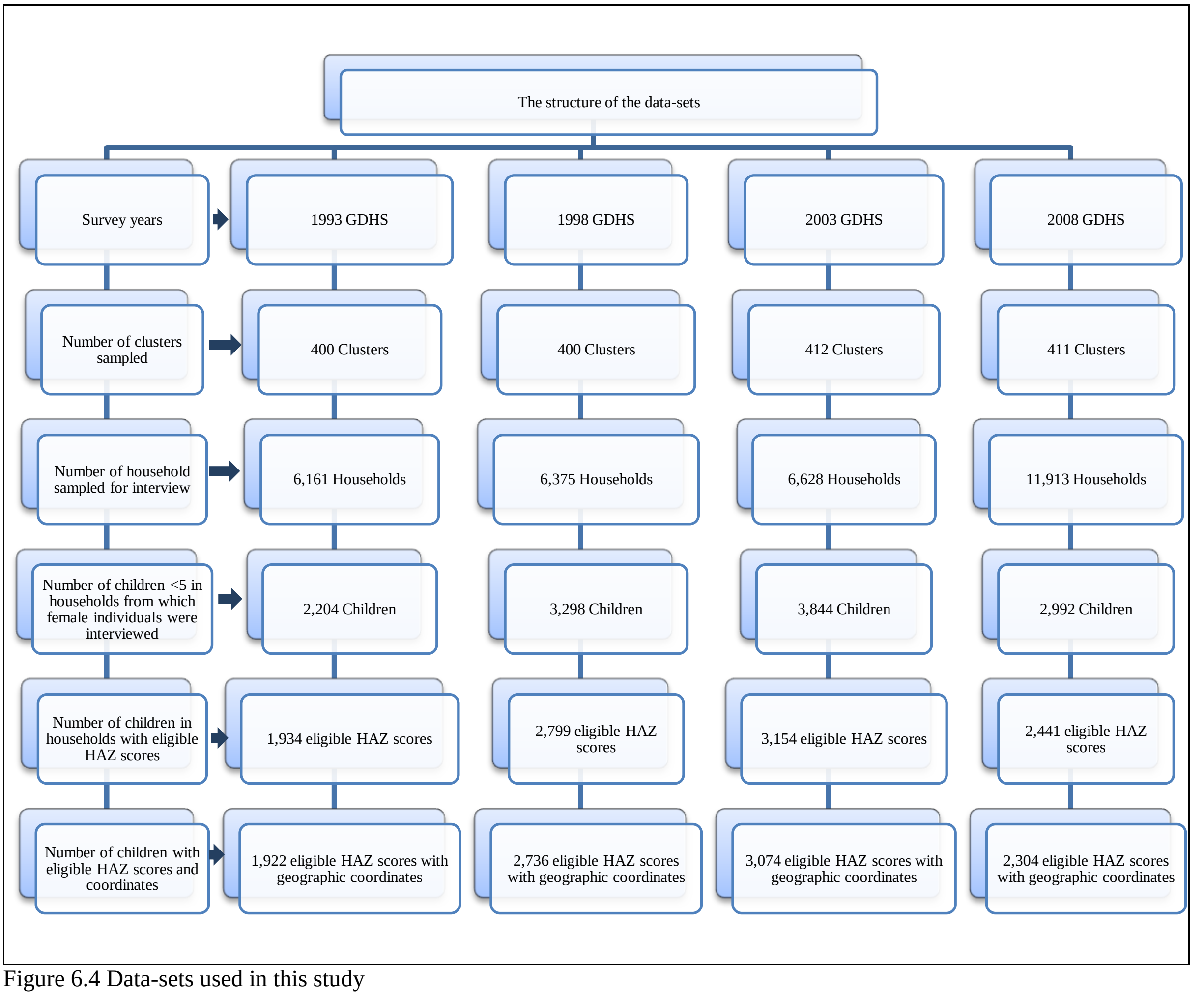


Figure 6.4 Data-sets used in this study

6.7 Annex 1: STROBE checklist

Title and abstract

1a. Indicate the study's design with a commonly used term in the title or the abstract

This has been done in abstract section. The study was a repeated cross sectional population based study.

1b. Provide in the abstract an informative and balanced summary of what was done and what was found

This has been done (page 112).

Introduction

2. Background/rationale. Explain the scientific background and rationale for the investigation being reported

This has been done. See pages 114-115.

3. Objectives. State specific objectives, including any prespecified hypotheses

This has been done. See second paragraphs at pages 114 and 115.

Methods

4. Study design. Present key elements of study design early in the paper

This has been done in the subsection "study population and design" (pages 115-116)

5. Setting. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection

This has been done in the subsection "study population and design" (pages 115-116)

6. Participants. Cross sectional study—Give the eligibility criteria, and the sources and methods of selection of participants

This has been done in the subsection "study population and design" (pages 115-116) and subsection "numbers of eligible participants" (page 118).

7. Variables. Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable

This has been done in the subsections "outcome variable" (page 117) and "explanatory variables" (page 117, and page 133 in Table 6.3)

8. Data source/measurements. For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group

This has been done in the subsections "study population and design" (pages 115-116), "outcome variable" (page 117), "explanatory variables" (page 117, and page 133 in Table 6.3) and "geographical data" (page 116-117).

9. Describe any efforts to address potential sources of bias

Source of possible bias have been included in the subsection "geographical data" (pages 116-117) detailing missingness in observations and have been stressed in study limitations (paragraph 4, page 130).

10. Study size. Explain how the study size was arrived at

This has been done in the subsections "study population and design" (pages 115-116) and "numbers of eligible participants" (pages 118).

11. Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why

This has been done in the subsection "model and methods" (page 119).

12. Statistical methods (a-e).

This has been done in the subsection "statistical analysis" (pages 118-124).

Results

13. (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. (b) Give reasons for non-participation at each stage(c) Consider use of a flow diagram

This has been done in the section “results” (page 124). Flow diagram has been included because it was a repeated cross sectional study (see Figure 6.4, page 133).

14. Descriptive data. (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Cohort study—Summarise follow-up time (eg average and total amount)

This has been done in the section “results” (see page 124 and Figure 6.4 at page 133). Follow-up time was not described because of the repeated cross sectional nature of the study.

15. *Outcome data. Cross sectional study—Report numbers of outcome events or summary measures*

Summary measure of outcomes has been discussed at each specific point.

16. *Main results. (a) Report the numbers of individuals at each stage of the study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram*

This has been included in results and tables. Flow diagram has been included because it was a repeated cross sectional study (see Figure 6.4, page 133).

17. *Other analyses. Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses*

Sensitivity analysis conducted (see pages 120-121, starting from last paragraph at page 120).

Discussion

18. *Key results. Summarise key results with reference to study objectives*

Key results have been summarised (see pages 128-129).

19. *Limitations. Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias*

Limitations have been discussed at page 130 in paragraph 4.

20. *Interpretation. Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence*

A cautious interpretation considering other studies has been done (see pages 128-131).

21. *Generalisability. Discuss the generalisability (external validity) of the study results*

Generalisability has been discussed at each specific point.

Other information

22. *Funding. Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based*

This has been included at page 113.

References

Adekanmbi, V. T., G. A. Kayode, and O. A. Uthman, (2013a), Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis: Maternal and Child Nutrition, v. 9, p. 244-59.

Adekanmbi, V. T., O. A. Uthman, and O. M. Mudasiru (2013b). Exploring variations in childhood stunting in Nigeria using league table, control chart and spatial analysis: BMC Public Health, v. 13, p. 361.

Aheto, J. M. K., T. J. Keegan, B. M. Taylor, and P. J. Diggle (2015). Childhood Malnutrition and Its Determinants among Under-Five Children in Ghana, Paediatric and Perinatal Epidemiology, p. 552-561.

Alom, J., M. A. Quddus, and M. A. Islam (2012). Nutritional status of under-five children in Bangladesh: a multilevel analysis: J Biosoc Sci, v. 44, p. 525-35.

Amugsi, D. A., M. B. Mittelmark, A. Lartey, D. J. Matanda, and H. B. Urke (2014). Influence of childcare practices on nutritional status of Ghanaian children: a regression analysis of the Ghana Demographic and Health Surveys: BMJ Open, v. 4, p. e005340.

Antwi, S. (2008). Malnutrition: missed opportunities for diagnosis: *Ghana Medical Journal*, v. 4**2**, p. 101-104.

Babatunde, R. O., F. I. Olagunju, S. B. Fakayode, and F. E. Sola-Ojo (2011). Prevalence and Determinants of Malnutrition among Under-five Children of Farming Households in Kwara State, Nigeria: Journal of Agricultural Science, v. 3.

Black, R. E., L. H. Allen, Z. A. Bhutta, L. E. Caulfield, M. de Onis, M. Ezzati, C. Mathers, J. Rivera, and M. a. C. U. S. Group (2008). Maternal and child undernutrition: global and regional exposures and health consequences: Lancet, v. 371, p. 243-60.

Bomela, N. J. (2009). Social, economic, health and environmental determinants of child nutritional status in three Central Asian Republics: Public Health Nutr, v. 12, p. 1871-7.

Brugha, R., and J. Kevany (1994). Determinants of nutrition status among children in the eastern region of Ghana: J Trop Pediatr, v. 40, p. 307-11.

de Onis, M., and M. Blössner (1997). WHO Global Database on Child Growth and Malnutrition, Geneva, WHO.

de Onis, M., and M. Blössner (2003). The World Health Organization Global Database on Child Growth and Malnutrition: methodology and applications: Int J Epidemiol, v. 32, p. 518-26.

Ezzati, M., A. D. Lopez, A. Rodgers, S. Vander Hoorn, C. J. Murray, and C. R. A. C. Group (2002). Selected major risk factors and global and regional burden of disease: Lancet, v. 360, p. 1347-60.

Feed the Future (2015). Ghana Country Profile. [Last accessed October, 2015] (http://www.feedthefuture.gov/country/ghana).

Food and Agriculture Organization (FAO) (2009). Nutrition Country Profiles-Ghana. [last accessed May 2015] (http://www.fao.org/ag/AGN/nutrition/gha_en.stm), Rome, FAO, United Nations.

Ghana Statistical Service (2008). Ghana Living Standards Survey: Report of the Fifth Round (GLSS 5), Ghana, GSS.

Ghana Statistical Service, Ghana Health Service, and ICF Macro (2009). Ghana Demographic and Health Survey 2008, Accra, Ghana., GSS, GHS, ICF Macro.
Ghana Statistical Service, Noguchi Memorial Institute for Medical Research, and ORC Macro (2004). Ghana Demographic and Health Survey 2003, Calverton, Maryland, GSS, NMIMR, and ORC Macro.
Ghana Statistical Service (GSS), and Macro International Inc (MI) (1994). Ghana Demographic and Health Survey 1993, Calverton, Maryland, GSS and MI.
Ghana Statistical Service (GSS), and Macro International Inc (MI) (1999). Ghana Demographic and Health Survey 1998, Calverton, Maryland, GSS and MI.
Kabubo-Mariara, J., G. K. Ndenge, and D. K. Mwabu (2009). Determinants of Children's Nutritional Status in Kenya: Evidence from Demographic and Health Surveys Journal of African Economies, v. 18, p. 363–387.
Kalman, R. E. (1960). A new approach to linear filtering and prediction problems Journal of Basic Engineering,Transactions ASME Series D, p. 35-45.
Kandala, N. B., T. P. Madungu, J. B. Emina, K. P. Nzita, and F. P. Cappuccio (2011). Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter?: BMC Public Health, v. 11, p. 261.
ORC Macro (2005). Nutrition of Young Children and Mothers in Ghana**:** Findings from the 2003 Ghana Demographic and Health Survey, Calverton, Maryland, USA, Africa Nutrition Chartbooks.
Pelletier DL, Frongillo EA, and Habicht JP (1993). Epidemiological evidence for a potentiating effect of malnutrition on child mortality: *American journal of Public Health*, v. **83**, p. 1133-1139.
Pelletier, D. L., and E. A. Frongillo (2003). Changes in child survival are strongly associated with changes in malnutrition in developing countries: J Nutr, v. 133, p. 107-19.
Pradhan, M., D. E. Sahn, and S. D. Younger (2003). Decomposing world health inequality: J Health Econ, v. 22, p. 271-93.
R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna, Austria.
Taylor, B. M. (2015). miscFuncs: Miscellaneous Useful Functions. R package version 1.2-7. http://CRAN.R-project.org/package=miscFuncs.
Taylor, B. M. (2010). Sequential Methodology and Applications in Sports-Rating. PhD Thesis. Lancaster University.
The World Bank (2006). Repositioning Nutrition as Central to Development: A Strategy for Large-Scale Action, Washington DC, The International Bank for Reconstruction and Development/The World Bank.
UNICEF, WHO, and The World Bank (2012). Joint child malnutrition estimates: Levels and trends. [last accessed 21/07/2015] (http://www.who.int/nutgrowthdb/estimates/en/), Geneva, WHO.
UNICEF-Ghana (2008). Issues facing children in Ghana., UNICEF. Available online at http://www.unicef.org/wcaro/Countries_1743.html.
Van de Poel, E., A. R. Hosseinpoor, C. Jehu-Appiah, J. Vega, and N. Speybroeck (2007). Malnutrition and the disproportional burden on the poor: the case of Ghana: Int J Equity Health, v. 6, p. 21.
WHO (1995). Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee: World Health Organ Tech Rep Ser, v. 854, p. 1-452.

WHO (2010). WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children., Geneva, WHO.
WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development, Geneva, WHO.

Chapter 7 Paper 4: Joint Modelling of Childhood Nutritional Outcomes among Under-Five Children in Ghana

This chapter is based on the following brief report:

This is a brief paper as a follow-up to paper 1 presented in this thesis. As a result, this paper is not meant to be a whole paper on its own to be submitted for publication and is considered to be an appendix to paper 1.

Abstract

Background: Childhood malnutrition is a global public health challenge and responsible for 40% of under-fives mortality in Ghana. In developing countries where childhood malnutrition remains prevalent, nutritional outcomes among children under-fives are commonly measured on weight-for-age (WAZ), height-for-age (HAZ) or weight-for-height (WHZ) Z-scores. Studies in developing countries investigating childhood nutrition and its predictors always analyse the data for each outcomes separately. The joint modelling of 2 or more nutritional outcomes therefore remains unstudied. This study investigates WAZ and HAZ jointly while allowing for the hierarchical nature of the data and accounting for important risk factors to help inform improved and targeted childhood nutrition interventions to minimise the risk.

Methods: The data collected in 2008 during the Ghana Demographic and Health Survey (GDHS) were used for the analysis. We analyse data on 2,083 children under-fives residing in 1,641 households, using multivariate response multilevel regression analysis.

Results: Our study found improvement in the level of accuracy and reliability in estimates for the joint models compared to the separate models. Our results showed that prolonged breastfeeding, multiple births, older ages, belonging to mothers with no national health insurance cover, and small size of the child at birth were inversely associated with both WAZ and HAZ scores. Mother's BMI was positively associated with both WAZ and HAZ scores while mother's age has positive association with HAZ only. Mother's education has positive association with WAZ only while diarrhoea episodes and absence of toilet facilities in households has inverse association with WAZ only. Belonging to poor households has inverse association with the HAZ only. Also, both residual individual and household-level effects for WAZ and HAZ are very strongly correlated and the impact of some statistically significant risk factors across WAZ and HAZ differ substantially. Substantial unexplained household-level variations in childhood nutritional outcomes for WAZ and HAZ scores were observed.

Conclusion: Our study has shown that studies into childhood nutrition among under-fives could benefit substantially from the joint modelling of nutritional outcomes to improve the level of accuracy and reliability in estimates and make inference about

residual correlations between nutritional outcomes on children and their households to better inform targeted childhood nutrition interventions to minimise the risk.

Keywords

Developing countries, Weight-for-age z-score, Height-for-age z-score, Child nutritional status, Multivariate response, Multilevel modelling, Joint modelling, Predictors of child nutritional status, Epidemiology, Biostatistics, Public health.

Funding

Funded by Faculty of Health and Medicine, Lancaster University, United Kingdom.

Except the funding of the study, the funder did not play any other role in this study including the brief report.

7.1 Introduction

In developing countries such as Ghana, malnutrition prevalence is high among children younger than five years and remains one of the leading causes of under-five deaths (Antwi, 2008; UNICEF-Ghana, 2008).

In such countries nutritional outcomes of children under the age of five are commonly measured on weight-for-age (WAZ), height-for-age (HAZ) and weight-for-height (WHZ) Z-scores (WHO, 1995). Height for age (HAZ) usually is taken as a measure of chronic malnutrition because poor nutrition over a child's lifetime will affect his or her long-term growth. Weight for height (WHZ) is usually taken as a measure of relatively recent food shortage, and weight for age (WAZ) is a combination measure. It is commonly found that in a population of under-five children such as Ghana's, the prevalence of these measures will vary, even when assessed on the same population (de Onis and Blössner, 1997). As they measure different types of malnutrition this is not unexpected.

Studies in developing countries investigating childhood malnutrition and its predictors mostly analyse the data for each outcome (WAZ, HAZ or WHZ scores) separately (Aheto et al., 2015; Alom et al., 2012; Babatunde et al., 2011; Bloss et al., 2004) and as a result, the joint modelling of 2 or more nutritional outcomes remains largely unstudied.

There are a number of reasons why it is interesting to investigate the malnutrition measure jointly. It is interesting to examine what factors are common drivers of both chronic and acute malnutrition. For example, modelling the nutritional outcomes jointly allows us to make inferences about the unmeasured factors that might have a common effect on the three measures of malnutrition (Snijders and Bosker, 2012). These outcomes when collected on the same individual are likely to be correlated. If in a statistical model the residual correlation from unmeasured factors is high then whatever this or these factors is or are likely to have a common effect on the three measures. This might help inform policy decisions.

Additionally, estimates obtained from joint modelling of two or more outcomes are expected to be more accurate and reliable, as indicated by their smaller standard errors compared to estimates from separate models (Snijders and Bosker, 2012). This could lead to positive impact on policy-making decisions. Joint modelling of outcomes is therefore preferable to separate analyses, both to make optimal use of the available information and to obtain unbiased estimates of the model parameters, especially when the main goal is to make inferences across outcomes.

This study aims to apply multivariate response multilevel models to jointly model WAZ and HAZ scores, to examine important risk factors for WAZ and HAZ scores and to investigate whether or not the effect of a risk factor on WAZ score is substantially different from its effect on HAZ score.

We also will examine whether or not after accounting for the risk factors in our model, the as-yet unidentified factors at the child or household level that determines WAZ and HAZ scores are the same and to estimate the variance explained by households in WAZ and HAZ scores among children.

Our goal is to improve the level of accuracy and reliability in estimates of the model parameters for WAZ and HAZ scores and to make optimal use of available data to help inform better and targeted childhood nutrition intervention strategies.

7.2 Methods

7.2.1 Study population

Data sourced from the 2008 Ghana Demographic and Health Survey (GDHS) were used for the analysis (Ghana Statistical Service et al., 2009). This study is a follow-up to previous research in which we used this data-set to investigate major risk factors for weight-for-age (WAZ), height-for-age (HAZ) and weight-for-height (WHZ) Z-scores in young children in Ghana using multilevel models to analyse WAZ, HAZ and WHZ separately (Aheto et al., 2015).

The data collection method is described elsewhere (Ghana Statistical Service et al., 2009), but briefly: the data collected from the Ghana Population and Housing Census conducted in 2000 were used as a sampling frame from which 412 clusters (communities) were selected nationwide for the 2008 GDHS. Anthropometric measurements as well as socio-economic and demographic characteristics were collected on 4,916 women aged 15-49 years and 4,568 men aged 15-59 residing in 11,778 selected households interviewed in the survey. Data on 2,992 children were extracted from the sampled women during the main survey. We excluded 756 children due to missing exposure data and we excluded a further 153 children because their Z-scores were outside the range of biologically plausible values as determined by World Health Organization (WHO) (WHO, 2010). The analysis in the present study was therefore based on 2,083 children younger than 5 years residing in 1,641 households with complete anthropometric measurements (outcomes) and information on their risk factors considered in this study from the women's data-set.

7.2.2 Outcome variables

The outcomes of interests for this analysis were weight-for-age or WAZ (a measure of underweight) and height-for-age or HAZ (a measure of stunting) scores. These were generated using the macro provided by WHO to obtain gender specific Z-scores (WHO, 2010).

7.2.3 Explanatory variables

To select potential risk factors for this study, we were guided by the literature on factors influencing nutritional outcomes of children under-five in developing countries as well as UNICEF's conceptual framework to analyse malnutrition in mothers and children which recognises the basic, underlying and immediate causes of malnutrition (Adekanmbi et al., 2013; Alom et al., 2012; Babatunde et al., 2011; Bomela, 2009; Kandala et al., 2011; UNICEF, 1990). In addition, we included other risk factors such as national health insurance status of the mother which might influence the nutritional outcomes of the children. Presented in Supporting Information Appendix S1 is the description of explanatory variables used in this analysis. A backward elimination method was used to arrive at the risk factors included in our final model.

7.2.4 Statistical analysis

The nutritional outcomes, WAZ and HAZ scores, are both measures of long term (chronic) malnutrition in the population. Our aim in the present study is to jointly model WAZ and HAZ scores to improve accuracy and reliability in estimates and to investigate the degree of residual correlation (similarities) between WAZ and HAZ on children and their households and to examine the effect of a risk factor across WAZ and

HAZ as well as examining differences in child nutritional outcomes across households in Ghana. To achieve this, we applied a three-level multivariate response multilevel regression model (Snijders and Bosker, 2012; Thum, 1997) presented in Equation (7.1) to analyse WAZ and HAZ scores on a continuous scale and used maximum likelihood to obtain parameter estimates in our model. Our model representing a three-level multivariate response multilevel model with WAZ and HAZ outcomes for a child i living in household j is of the form:

$$\begin{aligned} \mathrm{WAZ}_{ij} &= \beta^{(1)}\mathrm{X}_{ij} + \mathrm{h}_j^{(1)} + \mathrm{e}_{ij}^{(1)} \\ \mathrm{HAZ}_{ij} &= \beta^{(2)}\mathrm{X}_{ij} + \mathrm{h}_j^{(2)} + \mathrm{e}_{ij}^{(2)} \end{aligned} \tag{7.1}$$

$$\text{with} \quad \begin{pmatrix} \mathrm{h}_j^{(1)} \\ \mathrm{h}_j^{(2)} \end{pmatrix} \sim \mathrm{MVN}\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{h(1)} & \\ \sigma_{h(1,2)} & \sigma^2_{h(2)} \end{pmatrix}\right\}$$

$$\begin{pmatrix} \mathrm{e}_{ij}^{(1)} \\ \mathrm{e}_{ij}^{(2)} \end{pmatrix} \sim \mathrm{MVN}\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{e(1)} & \\ \sigma_{e(1,2)} & \sigma^2_{e(2)} \end{pmatrix}\right\}$$

where, WAZ_{ij} and HAZ_{ij} are the outcome variables namely weight-for-age and height-for-age Z-scores respectively for specific child *i* from the *j*th household; X_{ij} is the covariate that can be defined at the child or household levels; $\beta^{(1)}$ and $\beta^{(2)}$ are vector of regression coefficients for WAZ and HAZ respectively. The quantities $\mathrm{e}_{ij}^{(1)}$and $\mathrm{e}_{ij}^{(2)}$ are the residuals at the individual child level for WAZ and HAZ respectively and the quantities $\mathrm{h}_j^{(1)}$and $\mathrm{h}_j^{(2)}$ are the residuals at the household level for WAZ and HAZ respectively. In other words, $\mathrm{h}_j^{(1)}$ and $\mathrm{h}_j^{(2)}$ are the deviations from the overall mean for the *j*th household for WAZ and HAZ respectively and $\mathrm{e}_{ij}^{(1)}$ and $\mathrm{e}_{ij}^{(2)}$ are the deviations of the *i*th child from the mean of the *j*th household for WAZ and HAZ respectively. In addition, the quantities $\sigma^2_{h(1,2)}$ and $\sigma^2_{e(1,2)}$ are the covariance at the household and individual child-levels respectively for WAZ and HAZ scores.

The within-group variance component provided a good fit to our data among competing covariance structures. To examine the effect of a risk factor across WAZ and HAZ score, note that we only tested the effects of the significant risk factors common to both WAZ and HAZ scores in our models. We used the Stata package 'runmlwin' (Leckie and Charlton, 2013) to fit our models.

The population correlation coefficients (PCC) at the household (ρ_2) and child (ρ_1) levels are, respectively,

$$\rho_2 = \frac{\sigma_{u(1,2)}}{\sqrt{(\sigma^2_{u(1)} \times \sigma^2_{u(2)})}}, \quad and\ \rho_1 = \frac{\sigma_{\varepsilon(1,2)}}{\sqrt{(\sigma^2_{\varepsilon(1)} \times \sigma^2_{\varepsilon(2)})}}$$

Also, the intraclass correlation coefficient (ICC) at the household level for WAZ (ρ_{waz}) and HAZ (ρ_{haz}) are, respectively,

$\rho_{waz} = (\sigma^2_{u(1)})/(\sigma^2_{u(1)} + \sigma^2_{\varepsilon(1)})$ and $\rho_{haz} = (\sigma^2_{u(2)})/(\sigma^2_{u(2)} + \sigma^2_{\varepsilon(2)})$. In our study, the PCC measures the degree to which the residual correlations (similarities) between WAZ and HAZ scores depend on children and their households. Also, the ICC which coincides with the variance partition coefficient (VPC) measures the expected correlation between two children from the same household while the VPC measures the proportion of total variance which lies at the household level.

In a previous study (Aheto et al., 2015), we explored major risk factors for malnutrition and examined household level variations in childhood nutrition using separate multilevel models. Here we model the nutritional outcomes of children jointly using multivariate response multilevel models. Apart from exploring major risk factors for childhood nutrition and household level variations in childhood nutrition, we also investigate whether or not the effect of a risk factor on WAZ score is different from its effect on HAZ score and examine whether or not after accounting for the risk factors in our model, the as-yet unidentified factors at the child or household level that determines WAZ and HAZ scores are the same (similar).

It is important to mention that the present study also analysed all the three outcomes WAZ, HAZ and WHZ scores jointly but the inclusion of WHZ score led to the violation of the multivariate normality assumption underlying our model (see Figure 7.2 under Supporting Information Appendix S2). As a result, we excluded WHZ score in our analysis and considered only WAZ and HAZ scores in our final analysis that met the multivariate normality assumption (see Figure 7.1 under Supporting Information Appendix S2). For detailed description of the statistical method used in this analysis, see Chapter 3 section 3.1.2 of this thesis.

7.3 Results

7.3.1 Correlates for weight-for-age and height-for-age Z scores for multivariate response multilevel models

This study analysed data on 2,083 children, all having eligible Z-scores on both outcomes. Table 7.1 presents the results for the multivariate response multilevel model for weight-for-age (WAZ) versus height-for-age (HAZ) Z-scores. Regression coefficients are both positive and negative - predictor variable with negative coefficient suggests an adverse effect childhood nutrition status while a predictor variable with positive coefficient suggests a positive contribution to childhood nutrition status.

The results showed that risk factors such as prolonged breastfeeding, multiple births, older age, belonging to mothers with no national health insurance cover, and small size of the child at birth were statistically significantly, inversely associated with both WAZ and HAZ scores (i.e. showed an adverse effect on child nutritional status). Mother's

BMI was statistically significantly, positively associated with both WAZ and HAZ scores (i.e. indication of a positive contribution to child nutritional status) while mother's age has statistically significant, positive association with HAZ only.

Table 7.1 Parameter estimates for multivariate response multilevel models of weight-for-age (WAZ) versus height-for-age (HAZ) Z-scores.

	Parameter estimates	
	WAZ (Standard error)	HAZ (Standard error)
Child level characteristics		
Age in years	- 0.038* (0.0182)	-0.170** (0.0245)
Months of breastfeeding	-0.015**(0.0034)	-0.043**(0.0045)
Type of birth		
Single birth[a]		
Multiple birth	- 0.485** (0.1327)	- 0.470** (0.1724)
Diarrhoea		
No[a]		
Yes	- 0.117** (0.0435)	-
Fever		
No[a]		
Yes	- 0.063 (0.0435)	-
Size of child at birth		
Large/average[a]		
Small	- 0.309** (0.0606)	- 0.230** (0.0810)
Maternal/ household characteristics		
Mother's body mass index	0.043** (0.0057)	0.035** (0.0076)
Mothers' education (years)	0.010* (0.0048)	-
Household wealth status		
Rich or average[a]		
Poor	-	-0.131* (0.0544)
Mothers' current age	0.006 (0.0035)	0.020** (0.0044)
Type of toilet facility		
Flush or pit[a]		
No facility	-0.125** (0.0432)	-
Number of dead siblings		
None[a]		
One or more	0.062 (0.0457)	-
Mother has health insurance		
Yes[a]		
No	-0.094*(0.0470)	-0.184** (0.0614)

[a]Reference category
** Significant at 1% level
*Significant at 5% level

In addition, mother's education has statistically significant, positive association with WAZ only while diarrhoea episodes and absence of toilet facilities in households has statistically significant, inverse association with WAZ only. Belonging to poor households has statistically significant, inverse association with the HAZ only.

The results presented in Table 7.2 showed that the variance attributable to the residual household-level variation after accounting for child and household risk factors is 32% and 21% for WAZ and HAZ scores respectively. In addition, the residual correlation for individual and household-level effects for WAZ and HAZ were 0.52 and 0.78 respectively (see Table 7.2). The residual correlation between WAZ and HAZ scores on children or households measures whether or not, after accounting for the risk factors in our model, the as-yet unidentified factors at the child or household level that determines WAZ and HAZ scores are the same. Hence the higher (stronger) the residual correlation, the more similar these as-yet unidentified factors influencing both WAZ and HAZ scores will be and vice-versa.

Table 7.2 Variance and residual correlation estimates for the multivariate response multilevel models of weight-for-age versus height-for-age Z-scores.

	Variances		Intra-household correlation coefficients
	Child	Household	Explained variation (%)[a]
WAZ	0.65**	0.30**	32
HAZ	1.33**	0.36**	21

	Residual correlation
Child	0.52
Household	0.78

[a]Ratio of household-level variance to total variance multiplied by 100.
** Significant at 1% level.

In our analysis, we tested the hypothesis that the effect of a risk factor on WAZ score is the same as the effect of the same risk factor on HAZ score. Our tests showed that the impacts of child's age and breast-feeding duration on WAZ were statistically significantly different from the effect of each risk factor on HAZ score but no such differences exist for type of birth, size at birth, mother's body mass index and health insurance status (see Table 7.3).

Table 7.3 Test for equality of risk factor effect across WAZ and HAZ scores.

Risk factors	Degree of freedom	Chi-Square value	P-Value
Child level characteristics			
Age in years	1	40.22	0.00
Months of breastfeeding	1	52.51	0.00
Type of birth: Multiple birth	1	0.01	0.91
Size at birth: Small	1	1.36	0.24
Maternal/ household characteristics			
Mother's body mass index	1	1.65	0.20
Mother has health insurance: No	1	3.15	0.07

7.4 Discussion

This study is a follow-up to our previous study in which we investigated major risk factors for weight-for-age (WAZ), height-for-age (HAZ) and weight-for-height (WHZ) Z-scores using multilevel models to model the WAZ, HAZ and WHZ separately (Aheto et al., 2015). In the present study, we jointly modelled WAZ and HAZ scores and examined major risk factors for WAZ and HAZ scores. We also investigated whether or not the effect of a risk factor on WAZ score is substantially different from its effect on HAZ score and examined whether or not when accounting for the risk factors in our model, the factors at the child or household level that determines WAZ and HAZ scores which are not considered in our model are the same. We also estimate the variance explained by households in WAZ and HAZ scores among children.

The choice of multivariate response multilevel regression analysis in this study is appropriate because we have observations nested within children nested within households and we also observed strong empirical positive correlation between WAZ and HAZ scores in our data-set. Not recognising the correlation between WAZ and HAZ scores on children and households in the data-set may have resulted in overestimation of standard errors for the regression coefficients. This could lead to spurious non statistical significance and subsequent incorrect inference.

Our results showed that, generally, the standard errors associated with our estimates for the multivariate response multilevel models were smaller compared to those generated by the separate multilevel regression models presented in our previous paper (Aheto et al., 2015). This means that apart from allowing inferences about correlations and risk factors across WAZ and HAZ scores, the use of multivariate response multilevel models as opposed to the separate multilevel regression models led to an improvement in accuracy and reliability in estimates of our model parameters for WAZ and HAZ scores.

A consequence was that risk factors such as toilet facilities at homes and mother's health insurance status which were not statistically significant in the separate multilevel models for WAZ scores were statistically significant for the multivariate response multilevel model for WAZ scores using the same data-set and included risk factors used in the separate multilevel models. This is as a result of accounting for the residual correlation between WAZ and HAZ scores on children and the households in which they reside.

Furthermore, we observed that the residual child and household effects for WAZ and HAZ scores are very strongly correlated after adjusting for important risk factors such as child's age, duration of breastfeeding, type of birth (single/multiple), diarrhoea and fever episodes, maternal BMI, education, current age and health insurance status, household wealth status among others and that the correlation was stronger for the residual household effects (residual correlation = 0.78) than residual child effects (residual correlation = 0.52). This also suggests that after adjusting for risk factors in our model, it is the same as-yet unidentified factors at household level that influence both WAZ and HAZ scores. These factors could be household food availability, sanitation, socioeconomic as well as combined individual child characteristics like type of food consumed, duration of feeding and medical care not considered in our analysis. We also observed significant residual household variations of 32% and 21% in the nutritional outcomes of under-five children for WAZ and HAZ respectively. This suggests that nutritional outcomes of children varied across households in which they live and that we have not been able to capture adequately in our model all household factors that might explain reasons for variations in nutritional outcomes associated with households in which a child resides.

We also investigated whether the effect of a risk factor on WAZ score is substantially different from its effect on HAZ score. For example, we tested the null hypothesis that the effect of child's age on WAZ score was not statistically significantly different from its effect on HAZ score. We observed that for the risk factors which are common to both WAZ and HAZ scores, and which were statistically significant in our model for both WAZ and HAZ scores, only the impacts of child's age and breast-feeding duration on WAZ score were statistically significantly different from the effect of each of these risk factors on HAZ score. This means that the impact of child's age and breast-feeding duration across WAZ and HAZ scores varied substantially. This could help prioritise which risk factors to consider for urgent nutrition intervention purposes amidst scarce public health resources.

The strengths of this study include its population-based nature with nationwide coverage and good quality data on a number of child and household characteristics. The method we used, multivariate response multilevel modelling, allowed the joint modelling of WAZ and HAZ scores and identification of the degree to which the residual correlations (similarities) between WAZ and HAZ scores depend on children and their households as well as identification of household-level variation from presently unidentified factors. It also allowed us to test the effect of a given risk factor across WAZ and HAZ scores.

The limitations in this study include incomplete data on outcomes and exposures. Also, it is difficult to directly measure household wealth status in Ghana. Because of this we used an asset-based index, which is generally considered a good proxy for household wealth status in developing countries. In addition we did not have complete data for all children for all variables, for example only 19% of children in the data set have measurements on complementary feeding which is an important explanatory variable to consider and was excluded in the analysis in order not to reduce the sample size substantially which in turn could affect the power of the study.

Our study is the first in developing countries to jointly model nutritional outcomes measured on WAZ and HAZ scores through the application of multivariate response multilevel methods to improve accuracy and reliability of estimates and to provide additional but vital information on the degree to which the residual correlations (similarities) between nutritional outcomes for WAZ and HAZ scores depend on children and their households while simultaneously accounting for potential risk factors. We have also shown how valid conclusions can be drawn about the effect of a given risk factor across WAZ and HAZ scores using multivariate response multilevel models which are not possible should separate models be fitted to WAZ and HAZ scores.

Though we considered 2 nutritional outcomes in this study, this method can be extended to model more than 2 nutritional outcomes.

The likelihood of malnutrition in children in Ghana deteriorates with increases in children's' age, longer duration of breastfeeding, multiple births, experience of diarrhoeal episodes, small size at birth, household poverty, lack of toilet facility at homes and having a mother without health insurance cover. These findings are supported by previous studies (Adekanmbi et al., 2013; Kandala et al., 2011; Van de Poel et al., 2007; Van de Poel et al., 2008).

In this analysis we have modelled the effect of age on HAZ as being linear as in Babatunde et al. (2011), Madise and Mpoma (1997) and Kabubo-Mariara et al. (2009). This choice enables us to detect if there is an increase or a decrease in the risk of malnutrition with increasing age, it will not, however detect or describe non-linear effects. In order to detect non-linear effects, we would have to fit a more complex model, including for instance the effect of age on HAZ modelled as a spline function; we did not pursue this here, but further research in the area could explore this issue. The declining nutritional outcome observed among children who were breastfed for longer duration could be linked to inadequate complementary and exclusive breastfeeding practices among mothers in Ghana. This could also be linked to inadequate resources to provide healthy and adequate nutrition to the children and this study also showed that children from poor homes are more likely to have poor nutritional status as observed by other studies (Adekanmbi et al., 2013; Food and Agriculture Organization (FAO), 2009; Van de Poel et al., 2007).

The declining nutritional outcomes observed among children who experienced diarrhoeal episodes reflect wastage of food nutrients. It could also reflect poor sanitation in homes (Madise and Mpoma, 1997). This suggests that improved sanitation in homes could help mitigate episodes of diarrhoea thereby improving nutritional outcomes of children. Also, our finding that children from mothers without health insurance cover were more likely to have poor nutritional status compared to their counterparts from mothers with health insurance suggests that maternal health seeking behaviour and access to health is crucial in improving children nutrition and health. Children from homes without toilet facilities had increased risk of poor nutrition as toilet facility in home has been shown as a health variable contributing to improved sanitation and nutrition in children (Madise and Mpoma, 1997).

The nutritional outcomes of children improve with increase in mother's body mass index (BMI), years of education and age, which are supported by other studies (Adekanmbi et al., 2013; Alom et al., 2012; Van de Poel et al., 2007). The improved nutritional status observed among children from mothers with higher BMI in this study is because a mother's BMI is a reflection of her overall health and socioeconomic status so improved levels of mothers BMI might suggests improvement in maternal health and this should result in improved nutrition and health for their children (Adekanmbi et al., 2013).

Our study showed that improvement in the level of maternal education will improve child nutritional status as educated mothers will be more informed about good feeding and health care practices (types of food consumed and health seeking behaviour) for their children and themselves (Adekanmbi et al., 2013; Babatunde et al., 2011; Kabubo-Mariara et al., 2009; Kandala et al., 2011). Children from older mothers are less likely to suffer from poor nutrition because older mothers are less at risk of physiological immaturity during child birth (Heaton et al., 2005; Van de Poel et al., 2007). This suggests that encouraging adult motherhood could help improve child nutritional status.

Overall our study results add to the body of knowledge that show how household and child factors influence the likelihood of children in Ghana being malnourished. These findings can be applied to other, similar, developing countries. The effect on malnutrition of certain factors such as duration of breastfeeding has been quantified and may be of use in developing strategies at the population level to reduce the prevalence of malnutrition in vulnerable children.

7.5 Conclusion

We have shown that studies into childhood nutrition among children under-fives could benefit substantially from the joint modelling of nutritional outcomes by helping to improve the level of accuracy and reliability in estimates of the model parameters, to examine whether or not when accounting for the risk factors in our model, the factors at the child or household level that determines WAZ and HAZ scores which are not accounted for in our model are the same and to examine the effect of important risk factors across the outcomes. This could help improve the understanding of childhood nutrition to better inform targeted childhood nutrition intervention strategies. We suggest that researchers interested in analysing 2 or more under-fives nutritional outcomes should consider modelling those outcomes jointly as opposed to analysing those outcomes separately, especially in the presence of at least moderate empirical correlation between the outcomes.

7.6 Appendix

7.6.1 Supporting Information Appendix S1: Supplementary material related to the description of explanatory variables used in this analysis.

Table 7.4 Description of explanatory variables used in this analysis

Child characteristics
Age: Age in years and is continuous
Breastfeeding duration: breastfeeding duration in months and continuous
Type of birth: 1= multiple birth and 0= single birth
Diarrhoea episodes: 1= child had diarrhoea and 0= child did not have diarrhoea
Fever episodes: 1= child had fever and 0= child did not have fever
Size at birth: 1= small and 0= average/large
Maternal/household characteristics
Mother's education: mother's education in years and is continuous
Mother's body mass index (BMI): mother's BMI and is continuous
Mother's age: mother's age in years and is continuous
Mother's health insurance status: 1= mother did not have health insurance and 0= mother had health insurance
Household wealth status: 1= poor and 0= average/rich
Type of toilet facility in household: 1= no facility and 0= flush/pit facility
Number of dead siblings in household: 1= one or more and 0= none

7.6.2 Supporting Information Appendix S2: Information relating to our multivariate normality diagnostics for our multivariate response multilevel model for WAZ and HAZ, and WAZ, HAZ and WHZ scores.

Assessing multivariate normality assumption in household (Level 2) and child (Level 1) residuals for WAZ and HAZ scores.

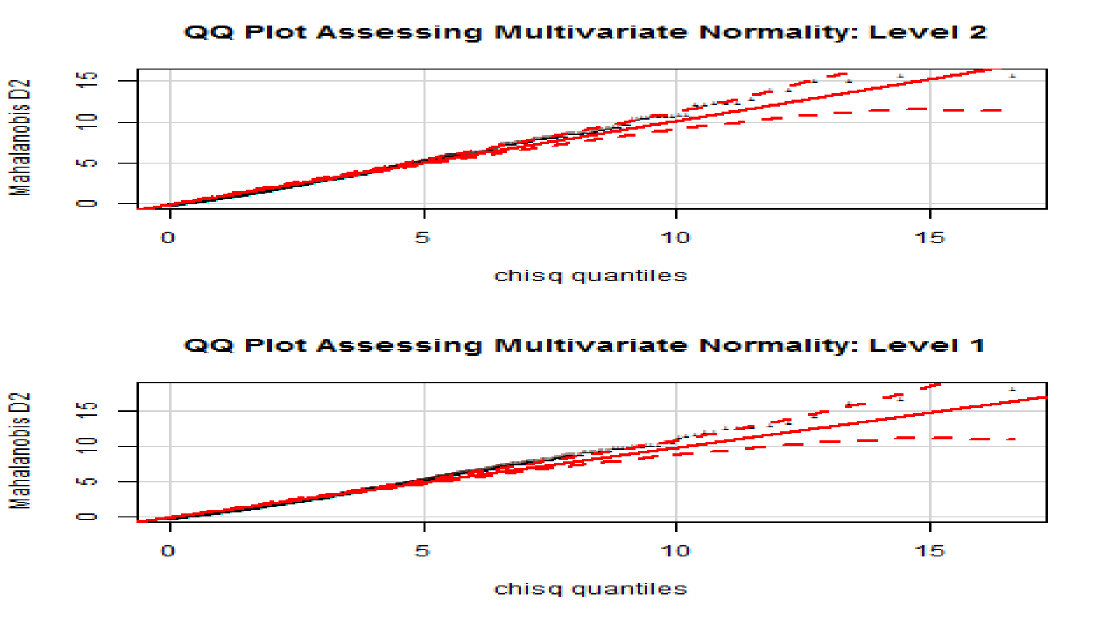


Figure 7.1 Multivariate normality plot for WAZ and HAZ scores.

Note: The multivariate normality assumption approximately passed when WAZ and HAZ scores were jointly modelled.

Assessing multivariate normality assumption in household (Level 2) and child-level (Level 1) residuals for WAZ, HAZ and WHZ scores.

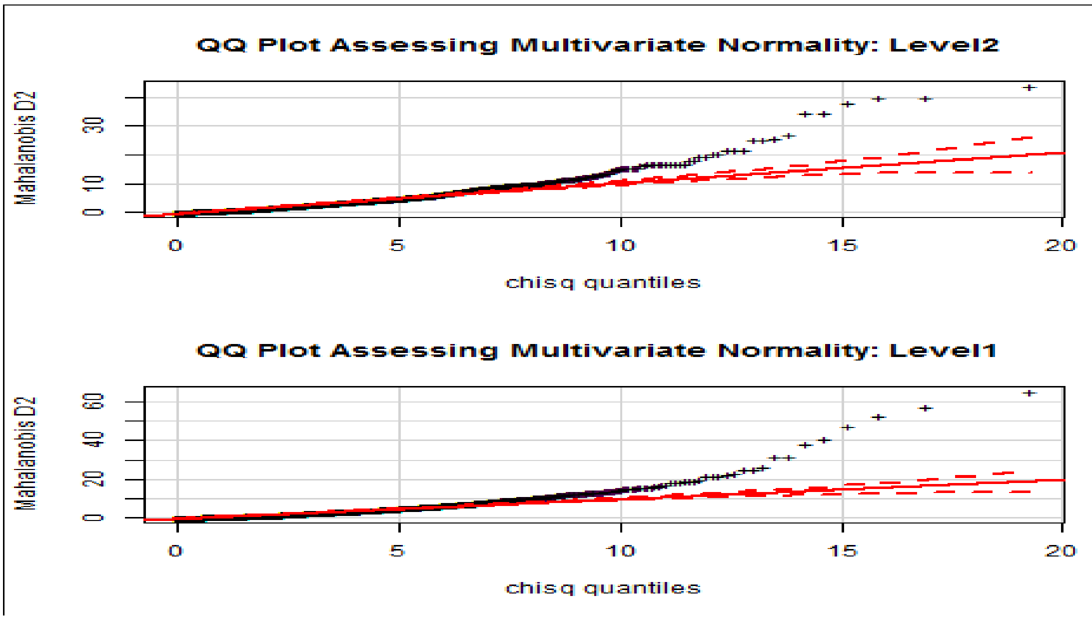


Figure 7.2 Multivariate normality plot for WAZ, HAZ and WHZ scores

Note: The multivariate normality assumption failed when WAZ, HAZ and WHZ scores were jointly modelled.

7.7 Annex 1: STROBE checklist

Title and abstract

1a. Indicate the study's design with a commonly used term in the title or the abstract
This has been done in abstract subsections. The study was a cross sectional population based study.

1b. Provide in the abstract an informative and balanced summary of what was done and what was found
This has been done (page 140).

Introduction

2. Background/rationale. Explain the scientific background and rationale for the investigation being reported
This has been done. See pages 142-143.

3. Objectives. State specific objectives, including any prespecified hypotheses
This has been done. See paragraphs 3, 4 and 5 at page 143.

Methods

4. Study design. Present key elements of study design early in the paper
This has been done in the subsection "study population" (page 144)

5. Setting. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
This has been done in the subsection "study population" (page 144)

6. Participants. Cross sectional study—Give the eligibility criteria, and the sources and methods of selection of participants
This has been done in the subsection "study population" (page 144)

7. Variables. Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
This has been done in the subsections "outcome variables" (page 145) and "explanatory variables" (page 145 and page 158 in Table 7.4)

8. Data source/measurements. For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Source of possible bias have been described in the subsection "study population" (page 144) detailing missingness in observations and have been stressed in study limitations (page 154, paragraph 2).

10. Study size. Explain how the study size was arrived at
This has been done in the subsection "study population" (page 144).

11. Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
This has been done in the subsection "statistical analysis" (see pages 145-146).

12. Statistical methods (a-e).
This has been done in the subsection "statistical analysis" (see pages 145-148).

Results

13. (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. (b) Give reasons for non-participation at each stage(c) Consider use of a flow diagram

This has been done in the section “results” (page 148). Flow diagram was not necessary because it was a one step cross sectional study.

14. Descriptive data. (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Cohort study—Summarise follow-up time (eg average and total amount)

This has been done (see page 148). Follow-up time was not described because of the cross sectional nature of the study.

15. *Outcome data. Cross sectional study—Report numbers of outcome events or summary measures*

Summary measure of outcomes has been discussed at each specific point.

16. *Main results. (a) Report the numbers of individuals at each stage of the study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram*

This has been included in results and tables. Flow diagram was not necessary because it was a onetime cross sectional study.

17. *Other analyses. Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses*

No other analysis has been done.

Discussion

18. *Key results. Summarise key results with reference to study objectives*

Key results have been summarised (see pages 152-153).

19. *Limitations. Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias*

Limitations have been discussed at page 154, paragraph 2.

20. *Interpretation. Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence*

A cautious interpretation considering other studies has been done (see pages 154-157).

21. *Generalisability. Discuss the generalisability (external validity) of the study results*

Generalisability has been discussed at each specific point.

Other information

22. *Funding. Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based*

This has been included at page 141.

References

Adekanmbi, V. T., G. A. Kayode, and O. A. Uthman (2013). Individual and contextual factors associated with childhood stunting in Nigeria: a multilevel analysis: Maternal and Child Nutrition, v. 9, p. 244-59.

Aheto, J. M. K., T. J. Keegan, B. M. Taylor, and P. J. Diggle (2015). Childhood Malnutrition and Its Determinants among Under-Five Children in Ghana, Paediatric and Perinatal Epidemiology, p. 552-561.

Alom, J., M. A. Quddus, and M. A. Islam (2012). Nutritional status of under-five children in Bangladesh: a multilevel analysis: J Biosoc Sci, v. 44, p. 525-35.

Antwi, S. (2008). Malnutrition: missed opportunities for diagnosis: *Ghana Medical Journal*, v. 4**2**, p. 101-104.

Babatunde, R. O., F. I. Olagunju, S. B. Fakayode, and F. E. Sola-Ojo (2011). Prevalence and Determinants of Malnutrition among Under-five Children of Farming Households in Kwara State, Nigeria: Journal of Agricultural Science, v. 3, p. 173-181.

Bloss, E., F. Wainaina, and R. Bailey (2004). Prevalence and Predictors of Underweight, Stunting, and Wasting among Children Aged 5 and Under in Western Kenya: Journal of Tropical Pediatrics, v. 50, p. 260-270.

Bomela, N. J. (2009). Social, economic, health and environmental determinants of child nutritional status in three Central Asian Republics: Public Health Nutr, v. 12, p. 1871-7.

de Onis, M., and M. Blössner (1997). WHO Global Database on Child Growth and Malnutrition, Geneva, WHO.

Food and Agriculture Organization (FAO) (2009). Nutrition Country Profiles-Ghana. [last accessed May 2015] (http://www.fao.org/ag/AGN/nutrition/gha_en.stm), Rome, FAO, United Nations.

Ghana Statistical Service, Ghana Health Service, and ICF Macro (2009). Ghana Demographic and Health Survey 2008, Accra, Ghana., GSS, GHS, ICF Macro.

Heaton, T. B., R. Forste, J. P. Hoffmann, and D. Flake (2005). Cross-national variation in family influences on child health: Soc Sci Med, v. 60, p. 97-108.

Kabubo-Mariara, J., G. K. Ndenge, and D. K. Mwabu (2009). Determinants of Children’s Nutritional Status in Kenya: Evidence from Demographic and Health Surveys Journal of African Economies, v. 18, p. 363–387.

Kandala, N. B., T. P. Madungu, J. B. Emina, K. P. Nzita, and F. P. Cappuccio (2011). Malnutrition among children under the age of five in the Democratic Republic of Congo (DRC): does geographic location matter?: BMC Public Health, v. 11, p. 261.

Leckie, G., and C. Charlton (2013). runmlwin - A Program to Run the MLwiN Multilevel Modelling Software from within Stata: Journal of Statistical Software, v. 52, p. 1-40.

Madise, N., and M. Mpoma (1997). Child malnutrition and feeding practices in Malawi: Food and Nutrition Bulletin, v. 18, p. 190-201.

Snijders, T. A. B., and R. J. Bosker (2012). Multilevel analysis: an introduction to basic and advanced multilevel modeling: London, SAGE.

Thum, Y. M. (1997). Hierarchical Linear Models for Multivariate Outcomes, Journal of Educational and Behavioral Statistics, p. 77-108.

UNICEF (1990). Strategy for improved nutrition of children and women in developing countries: a UNICEF policy review, New York, UNICEF, p. 22.

UNICEF-Ghana (2008). Issues facing children in Ghana. [Cited 21/07/2015]. Available online at http://www.unicef.org/wcaro/Countries_1743.html, UNICEF.

Van de Poel, E., A. R. Hosseinpoor, C. Jehu-Appiah, J. Vega, and N. Speybroeck (2007). Malnutrition and the disproportional burden on the poor: the case of Ghana: Int J Equity Health, v. 6, p. 21.

Van de Poel, E., A. R. Hosseinpoor, N. Speybroeck, T. Van Ourti, and J. Vega (2008). Socioeconomic inequality in malnutrition in developing countries: Bull World Health Organ, v. 86, p. 282-91.

WHO (1995). Physical status: the use and interpretation of anthropometry. Report of a WHO Expert Committee: World Health Organ Tech Rep Ser, v. 854, p. 1-452.

WHO (2010). WHO Anthro for personal computers, version 3.2.2, 2011: Software for assessing growth and development of the world's children., Geneva, WHO.

Chapter 8 General discussion and future studies

The applications of appropriate and advanced statistical methods in this thesis are motivated by a real-life chronic public health childhood malnutrition problem which is one of the global major public health challenges, especially in developing countries.

We presented the results on major risk factors for malnutrition and unexplained household-level variations in nutritional outcomes of children in Chapter 4. Such a study is crucial in terms of identifying important factors influencing nutritional outcomes of children and whether or not there exist differences in nutritional outcomes of children across households to inform appropriate policies and intervention strategies aimed at improving childhood nutrition and health.

In Chapter 5, we presented the results on the spatial analysis. This study investigated whether or not childhood malnutrition vary over place, identified communities at higher risk of childhood malnutrition and examined climatic and environmental factors that might affect the spatial patterns. This study is first of its kind and very important in terms of improving the understanding of how childhood malnutrition situations vary across Ghana and provided deeper insight into factors influencing it. This study also provided the opportunity for the higher risk communities to be identified and targeted for public health interventions, thereby making effective and optimal use of limited public health resources available in the country; the limited resources will be spent on those communities that needed it most unlike the current nutrition policies and interventions conducted on mass basis which are less or not supported with sound data and analysis and not directed to those that needed it most. Thus, the limited public health resources allocated for nutrition interventions should be channelled to the higher risk communities because they needed such interventions most. This thesis developed risk maps and identified priority areas for implementation of an intervention that might be known to be effective but is too expensive to implement universally, the first of its kind in the country.

In Chapter 6, we explored spatial and temporal patterns in the risk of childhood chronic malnutrition and provided forecast maps for the risk of childhood chronic malnutrition. The importance of this study is based on the fact that it will enable pre-emptive childhood malnutrition interventions as well as promoting prioritisation, effective and sustainable public health policies in the country. This study also helped identify the progress made so far in Ghana to reduce childhood malnutrition overtime and established that substantial differences exist in the malnutrition prevalence from one geographical location to another over the study time periods in the country. Childhood malnutrition prevalence, trend and its spatial distribution analysis are crucial for public health planning because it can help support carefully targeted interventions aimed at reducing childhood malnutrition amidst scarce public health resources.

In Chapter 7, we investigated the degree to which the residual correlations between nutritional outcomes measured on weight-for-age (WAZ- a measure of underweight) and height-for-age (HAZ - a measure of stunting) Z-scores depend on the children and their households which permitted the determination of whether or not after accounting for the risk factors in our model, the as-yet factors at the child or household level that influence both WAZ and HAZ scores are the same. We observed strong residual correlation between WAZ and HAZ scores on households. It is worth noting that this study is the first to apply multivariate response multilevel models to jointly model WAZ and HAZ among under-five children in Ghana. This could help inform policy decisions since any intervention that could investigate and address such as-yet unidentified factors influencing WAZ score can equally do for HAZ score. We also examined the major risk factors for under-fives nutrition as well as a specific effect of a risk factor across these outcomes. We also examine the variance explained by households in WAZ and HAZ scores among children. Such a study is crucial in terms of obtaining unbiased estimates of the model parameters for correct inferences leading to improved understanding of childhood nutrition and this should result in optimal public health policies and interventions. This study can also be a source of encouragement for other researchers interested in analysing 2 or more nutritional outcomes of children to consider modelling such outcomes jointly rather than separately.

Overall our study results add to the body of knowledge that show how child, household and environmental factors influence the likelihood of children in Ghana being malnourished and provided sound basis for child nutrition and health policies and interventions in the country. This study will enable limited public health resources allocated to child nutrition policies and interventions in Ghana to be spent on those communities that needed it most unlike the current nutrition policies and interventions conducted on mass basis which are less or not supported with sound data and analysis and not targeted to those at more risk of malnutrition.

The findings in this thesis can be applied to other, similar, developing countries. The effect on malnutrition of certain factors such as child diarrhoea episodes, maternal health insurance cover, maternal education and body mass index, toilet facilities, vegetation index, elevation, and the spatial and spatio-temporal distribution in the risk of child malnutrition over Ghana has been quantified and can be used in developing strategies at the population level to minimise the prevalence of malnutrition in vulnerable children, especially those from high risk communities.

The findings from this thesis have key policy and intervention implications for improving childhood nutrition and health in Ghana and more widely. The government of Ghana and non-governmental organizations should strive to provide free health care services for pregnant women, mothers, and children under five years of age. Public health measures should also build on the work already done to implement WHO recommendations on exclusive breast feeding up to 6 months of age. Also, the need to feed infants with nutritionally adequate and safe complementary foods from six months of age together with continued breastfeeding up to age two or beyond must be emphasized. Beyond economic redistribution, the government can provide subsidies for children's food and assistance for households in installing sanitary toilets in their homes to help improve the overall health and nutrition of young children. To increase levels of women's education more generally, governments such as Ghana's could make basic and secondary education more accessible, compulsory and affordable, if not completely free.

Furthermore, greater consideration should be given to the high-risk areas identified. Policy and intervention programmes targeting the higher-risk areas identified could be designed to include specific policies on childhood nutrition and health. There is a need to institute pragmatic and actionable strategies backed by a strong political will to develop and support the regions of Ghana to become major food baskets and growth poles through adequate exploitation of the abundant rich natural resources available in these regions. Community-based nutrition surveillance systems in the form of nutrition evaluation and monitoring programs should be introduced and used to monitor the nutritional outcomes and health of children under-five years of age regularly, so as to provide regular up-to-date nutrition information on the under-five population to aid timely and relevant nutrition intervention strategies. Issues relating to environmental and climatic conditions should also contribute to the formulation of public health nutrition intervention strategies.

The results presented in Chapters 4 and 7 showed that we have not been able to capture adequately in our model all household factors that might explain reasons for variations in nutritional outcomes associated with households in which children reside. Further study is therefore warranted to investigate other factors not considered in our analysis which might be responsible for the unexplained household-level variations in nutritional outcomes of children observed in this study. Also, the finding that longer duration of breast feeding was associated with more risk of malnutrition warrant further studies in the form of longitudinal analysis to establish its association. We believe that the relationship between extended breastfeeding and increased likelihood of malnutrition is not causal - that is it confounded by an absence of complimentary foods rather than an adequacy of breast feeding. This finding should therefore be interpreted with greatest caution.

In Chapters 5 and 6, substantial spatial variations in the risk of childhood malnutrition were observed. We therefore advocate for further studies to examine geographical, climatic or environmental factors not considered in our analysis which might help account for the unexplained spatial variations in the risk of childhood malnutrition observed in this study.

Furthermore, the forecast maps presented in Chapter 6 are for the year 2013 because at the time of writing this thesis, the 2008 Ghana Demographic and Health Survey is the most recent survey data-set available to the public. To illustrate the performance of our fitted spatio-temporal model, we therefore used the GDHS for the years 1993, 1998, 2003 and 2008 to forecast the probabilities of stunting for the year 2013. We intend to apply our current spatio-temporal modelling techniques to update our model and to provide forecasts for the year 2019, as soon as the 2014 GDHS data-set is made available to the public.

Curriculum Vitae

(Restricted to January 2013 – December, 2015: the period of my candidature for a research degree at Lancaster University)

Personal Details

Name : Justice Moses Kwaku Aheto
Gender : Male
Date of Birth : 25 February 1981
Office Address : Lancaster Medical School, Faculty of Health and Medicine, Lancaster University, Bailrigg, Lancaster LA1 4YB, United Kingdom
E-mail : j.aheto@lancaster.ac.uk
Web pages : http://www.lancaster.ac.uk/fhm/about-us/people/justice-aheto
: http://lancaster.academia.edu/JusticeMosesKAheto

Publications

1. **Aheto JMK**, Keegan TJ, Taylor BM, Diggle PJ (2015). Childhood Malnutrition and Its Determinants among Under-Five Children in Ghana. *Paediatric and Perinatal Epidemiology*; **29**:552-561.

2. **Aheto JMK**, Taylor BM, Keegan TJ, Diggle PJ (2015). Modelling Spatial Variations in the Risk of Childhood Malnutrition among Under-Five Children in Ghana. Under review in *Statistics in Medicine*.

3. **Aheto JMK**, Taylor BM, Keegan TJ, Diggle PJ (2015). Modelling and Forecasting Spatiotemporal Variations in Risk of Childhood Malnutrition among Under-Five Children in Ghana. Under review in *Spatial and Spatio-Temporal Epidemiology*.

Oral Presentation

1. Aheto JMK (2015). Analysing malnutrition prevalence and its determinants among under-five children in Ghana: Multilevel Methods. The 2015 Nutrition and Nurture in Infancy and Childhood: Bio-Cultural Perspective Conference, University of Central Lancashire, June 10-12. Cumbia, UK.

Awards

- TakeAIM Prize Winner (November 2015). The contest challenges researchers in mathematical sciences to describe their research in 250 words for a general audience, and is open to students from all European universities. My entry was based on the work in this thesis and was singled out as one of the best for its potential impact in the study of child malnutrition. The contest was organised by Smith Institute for Industrial Mathematics and System Engineering, London, UK. (£100)

- Travel grant from William Ritchie Travel Fund here at Lancaster University to attend and to present my paper titled "Analysing Malnutrition Prevalence and its Determinants among Under-Five Children in Ghana: Multilevel Methods" at 'Nutrition and Nurture in Infancy and Childhood: Bio-Cultural Perspectives Conference' organised by University of Central Lancashire. 10-12 June 2015, Cumbia, UK. (£300)

- Travel grant from the Faculty of Health and Medicine to attend R programming course at Newcastle University held in September 2014, Newcastle, UK. (£460)

- PhD Studentship by Faculty of Health and Medicine bursary funding for Doctoral Research at the overseas level covering tuition fees and stipend for 3 years.

Professional Activities

- **Associate Lecturer– Department of Mathematics and Statistics, Lancaster University (2013-2015):** teaching Medical Statistics (MATH335), Likelihood Inference (MATH330), Likelihood Inference (MATH551), Bayesian Inference (MATH331), Topics in Modern Statistics (MATH334), Project Skills (MATH390), Probability and Measure (MATH313), Integration (MATH314), Geometry of Curves and Surfaces (MATH329), Probability (MATH230), Probability (MATH104), Minor Course in Mathematics (MATH271), Minor Course in Mathematics (MATH272), Statistics (MATH105), Matrix Methods (MATH103), Integration (MATH102), Calculus (MATH101) and Application of Statistical Packages in Data Analysis, Data Analysis, Interpretation in Biomedical and Life Sciences (BIOL421). I am also responsible for invigilating end of module examinations for the department.

- **Member – Faculty and Health and Medicine Research Committee (2013-2015):** Responsible for bringing issues affecting postgraduate research students in the Faculty to the attention of the Committee for consideration as well as discussing all matters relating to research and consultancy in the Faculty, for the development of policy recommendations, and for the implementation of policy decisions of the Policy and Resources Committee. In addition, I have provided free consultancy to the Committee in which I carried out statistical analysis on the data collected by the Faculty and prepared reports on the reasons why potential students decided to come to Lancaster for their studies, especially to the Faculty of Health and Medicine and to determine who the University/Faculty's competitors (other universities) are and the courses they offer, both in the UK and globally.

- **Examination Invigilator, Lancaster University (2013 - 2015):** Responsible for invigilating end of year examinations for the university.

- **Student Registration Helper, Lancaster University (2013 - 2015):** Part of the team responsible for registering fresh students for the university.

PhD Training and Development

- Environmental Epidemiology (Spatial Analysis), Methods for Missing Data Analysis, Multilevel Models, Structural Equation Modelling, Secondary Data Analysis and Data Analysis with SPSSII. Department of Mathematics & Statistics, Lancaster University.

- R Programming Course. School of Mathematics & Statistics, September 2014. Newcastle University.

- Mathematics, Statistic and Operations Research: A workshop for Postgraduate Students who teach Mathematics, Statistics and Operations Research by Higher Education Academy STEM, UK, October 2013. Held at Lancaster University.

- Information Security course. ISS, September 2015. Lancaster University.

Professional Association Membership

- The Royal Statistical Society
  - Young Statistician Section

- The International Biometric Society
  - British and Irish Region