

https://dspace.org/

Spatial modelling of under five child mortality and associated factors in Zimbabwe using 2011 and 2015 demographic and health survey data

Sibanda, Morelearnings

2021
https://hdl.handle.net/10539/33614

Downloaded from a DSpace repository. https://dspace.org/

SPATIAL MODELLING OF UNDER FIVE CHILD MORTALITY AND ASSOCIATED FACTORS IN ZIMBABWE USING 2011 AND 2015 DEMOGRAPHIC AND HEALTH SURVEY DATA



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Morelearnings Sibanda (1620303)

A Research Report Submitted to The Faculty of Health Sciences, University of the Witwatersrand in Partial Fulfillment of the Requirements for the Degree Master of Science in Epidemiology and Biostatistics in the Field of Biostatistics

Johannesburg, October 2021

Approved for the Degree

(Master of Science in Epidemiology and Biostatistics)

Supervisor: Prof Ngianga Bakwin Kandala

Supervisor: Dr Zvifadzo Matsena Zingoni

Declaration

I, Morelearnings Sibanda, student number 1620303 declare that this research project is my work. It is being submitted in partial fulfillment of the degree of Master of Science in Epidemiology and Biostatistics in the field of Biostatistics at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.



Morelearnings Sibanda

October 2021

Dedication

Dedicated this dissertation to my loving parents and siblings who believed in me and supported me all the way.

Abstract

Background: Under-five mortality (U5M) is a global public health concern as it is a measure of a country's overall performance. Zimbabwe failed to achieve millennium development goal 4 (MGD) to reduce U5M by two thirds in 2015 and as we target to achieve sustainable development goal 3 (SDG) which aims to reduce U5M to 25 deaths in 1000 live births in 2030, there is need to review the progress during the last two surveys (2011 and 2015) of the MDG period and identify risk factors of U5M. Children from the same geographical region are affected by similar factors, some of which are unmeasured and accounting for these factors through the incorporation of spatial effects produces more informative results. Therefore, the aim of this study was to determine and compare the spatial distribution of U5M while adjusting for possible risk factors in Zimbabwe at the district level between the two surveys.

Methods: The current study was based on secondary data analysis of data collected in the 2011 and 2015 Zimbabwe demographic and health surveys (ZDHS). The study population consisted of under-five children aged 0-59 months (both alive and dead). To describe the data, we used frequencies and percentages for categorical variables while means and standard deviations were used for continuous data. Univariate analysis of survival experiences for categorical variables was visualised using non-parametric Kaplan-Meier (KM) plots. Bayesian spatial survival models based on Markov chain Monte Carlo (MCMC) simulation techniques were fit to determine the spatial distribution of U5M at the district level for both surveys.

Cox proportional hazards (Cox-PH) and accelerated failure time (AFT) models were the survival analysis models that were fit and they were adjusted for spatial frailty and possible risk factors of U5M. These models were run in OpenBugs and R statistical software and the best fit model was selected based on the lowest deviance information criterion (DIC) value. For visualisation of spatial heterogeneity of U5M at the district level, maps were plotted in both software for the two surveys.

Results: A total of 5,563 participants were considered in the 2011 survey and 6132 in the 2015 survey. After accounting for sampling weights, we found that 6.93% and 5.65% of the children died before five years in the 2011 and 2015 surveys respectively. Findings indicate that the U5M rates were significantly lower in 2015 compared to 2011. Relative to their respective reference categories which were singleton birth, being a boy child, average size at birth, being married, mother's non-use of contraceptives and Apostolic religion, the following results were obtained for U5M factors from the Cox-PH models. Multiple births (posterior HR (PHR) = 3.84; 95% CI: 2.56 - 5.50), being a girl child (PHR = 0.74; 95% CI: 0.60 - 0.90), the small size of child at birth (PHR = 1.73; 95% CI: 1.26 - 2.25), mother's single marital status (PHR = 0.32; 95% CI: 0.083 - 0.74), mother's contraceptive use (PHR = 0.53; 95% CI: 0.42 - 0.66), Pentecostal religion (PHR = 0.72; 95% CI: 0.51 - 0.94) and rural residence (PHR = 0.70; 95% CI: 0.48 - 1.00) were significantly associated with U5M in the 2011 survey. In 2015, multiple births (PHR = 4.52; 95% CI: 3.11 - 6.28), being a girl child (PHR = 0.79; 95% CI: 0.63 - 0.96), the small size of child at birth (PHR = 1.81; 95% CI: 1.32 - 2.37), the large size of child at birth (PHR = 1.47; 95% CI: 1.15 - 1.85), mother's single marital status (PHR = 0.55; 95% CI: 0.24-0.99) , mother's contraceptive use (PHR = 0.51; 95% CI: 0.41 - 0.64), Pentecostal religion (PHR = 0.46; 95% CI: 0.24- 0.78) and other religions (PHR = 0.61; 95% CI: 0.42-0.84) were significantly associated with U5M. Maps for exceedance probabilities suggested spatial clustering of U5M hotspots. Regions with higher chances of U5M hazards

and low survival times were clustered in the northern, middle and eastern parts of Zimbabwe, while regions with lower chances of hazards were clustered towards the southern and western parts of Zimbabwe for both survey points.

Conclusion: Information on the geographical distribution of U5M across the country is important in understanding how frailty due to unmeasured effects at the district level can affect child survival. There were geographical variations in U5M across the country and we identified districts from the northern, central and eastern parts to be hotspots that persisted between the two surveys. This points towards the need for implementation of strategies and targeted interventions at the subnational level to guide resource allocation to fast-track progress in reducing U5M. Even though there was significant progress in decreasing U5M between 2011 and 2015, the mortality rates remained relatively high and if Zimbabwe is to achieve SDG 3 by 2030, then fundamental steps must be taken towards tackling the U5M risk factors.

Acknowledgements

I would like to express my sincere appreciation to my supervisors Prof N B Kandala and Dr Z Matsena Zingoni for their guidance and mentorship during my research journey. I would also like to extend my gratitude to the Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) for funding my masters studies and making my dream possible. Special thanks to fellow Biostatisticians, Wits school of public health lecturers and staff for their assistance and support during my time at Wits.

Table of Contents

Declaration . . . . . **i**

Dedication . . . . . **ii**

Abstract . . . . . **iii**

Acknowledgements . . . . . **vi**

List of Figures . . . . . **xi**

List of Tables . . . . . **xii**

Nomenclature . . . . . **xiii**

Chapter 1: INTRODUCTION **1**

1.1 Background 1

1.2 Problem Statement 3

1.3 Justification 4

1.4 Conceptual Framework 6

1.5 Literature Review 7

1.5.1 Frailty in Survival Analysis 11

1.5.2 Bayesian Spatial Survival 11

1.6 Research Question 13

1.7 Objectives . . . 13

Chapter 2: METHODS . . . 14

2.1 Introduction . . . 14

2.2 Study Design/Site . . . 14

2.3 Data Sources . . . 16

2.4 Data Quality . . . 18

2.5 Current Study Population . . . 18

2.6 Data Management . . . 19

2.6.1 Data Extraction . . . 19

2.6.2 Data Preparation . . . 20

2.6.3 Missing Data . . . 20

2.7 The MICE method . . . 21

2.8 Spatial Data . . . 21

2.9 Statistical Analysis . . . 22

2.10 Kaplan-Meier Model . . . 24

2.11 Cox Proportional Hazards Model Specification . . . 24

2.12 Accelerated Failure Time model (AFT) Model Specification . . . 32

2.13 Model Fitting in R and OpenBugs . . . 37

2.14 Model Selection . . . 37

2.15 Model Assessment and Goodness of Fit . . . 38

2.16 Ethical Considerations . . . 38

Chapter 3: RESULTS . . . 40

3.1 Introduction . . . 40

3.2 Descriptive Statistics . . . 41

3.3 Mortality Rate: Standardised Mortality Ratio and KM Curves . . . 44

3.3.1 Standardised Mortality Ratio (SMR) . . . 44

3.3.2 Mortality Rate and KM Curves . . . 44
3.4 2011 Regression Models . . . 49
3.5 2015 Regression Models . . . 53
3.6 Spatial Distribution . . . 57
3.6.1 2011 Spatial Distribution . . . 57
3.6.2 2015 Spatial Distribution . . . 59

Chapter 4: DISCUSSION and CONCLUSION . . . 61
4.1 Introduction . . . 61
4.2 U5M Risk Factors . . . 62
4.3 Spatial Distribution of U5M . . . 65
4.4 Comparison between 2011 and 2015 Findings . . . 67
4.5 Modelling . . . 68
4.6 Strengths and Limitations . . . 68
4.6.1 Study Design . . . 68
4.6.2 Modelling Computational Aspects . . . 69
4.7 Conclusion . . . 70

References . . . 71

Appendix A: Plagiarism Declaration . . . 90

Appendix B: Missing Data . . . 91

Appendix C: Model Specification in R . . . 93

Appendix D: Model Selection . . . 98

Appendix E: Model Codes . . . 100
OpenBugs Model Codes . . . 100
R Model Cox-PH and AFT Codes . . . 102

Appendix F: Diagnostics Plots **103**

OpenBugs Diagnostics Plots 103

R Diagnostics Plots 105

Appendix G: Ethics Certificate **106**

Appendix H: Turnitin Report **107**

List of Figures

1.1 Mosley and Chen framework of child survival 6

1.2 Conceptual framework of under-five mortality 7

2.1 Map showing Zimbabwe provinces labelled by names 15

2.2 Map showing Zimbabwe districts labelled by names 16

2.3 Enumeration (cluster) points for 2011 (left) and 2015(right) ZDHS. . 17

3.1 Maps showing the standardised U5M ratio of the districts in Zimbabwe in 2011 (left) and 2015(right) ZDHS. 44

3.2 The overall Kaplan-Meier curve for the U5M rates for 2011 and 2015. 45

3.3 Kaplan-Meier curves for child survival in Zimbabwe by different characteristics. 49

3.4 2011 structured (left) and unstructured (right) spatial survival at district level from OpenBugs (top) and R (bottom). 57

3.5 2011 structured (left) and unstructured (right) effects maps of exceedance probability (PTR $<$ 1.0) at district level. 58

3.6 2015 structured (left) and unstructured (right) spatial survival at district level from OpenBugs (top) and R (bottom). 59

3.7 2015 structured (left) and unstructured (right) effects maps of exceedance probability (PHR $>$ 1.0) at district level. 60

List of Tables

3.1 Baseline Characteristics of study participants, for 2011 and 2015 data, stratified by mortality. 42

3.2 Mortality rate per 1000 person-years and log-rank test of under-five children in Zimbabwe for 2011 and 2015 ZDHS data. 46

3.3 Adjusted non-spatial and spatial Cox-PH and AFT models for 2011 ZDHS data. 50

3.4 Adjusted non-spatial and spatial Cox-PH and AFT models for 2015 ZDHS data. 54

Nomenclature

Definition of terms

Term	**Definition**
Infant Mortality	Death of a child before the first birthday
Child Mortality	Death of a child after the first birthday but before the fifth birthday

Abbreviations

Abbreviation	**Phrase**
AFT	Accelerated Failure Time
CAR	Conditional Autoregressive prior
EA	Enumeration Area
KM	Kaplan - Meier
MICE	Multivariate Imputation by Chained Equations
MLE	Maximum Likelihood Estimate
PH	Proportional Hazards
PO	Proportional Odds
SDG	Sustainable Development Goals.
SSA	Sub-Saharan Africa
U5M	Under-five Mortality
UN	United Nations
ZDHS	Zimbabwe Demographic and Health Survey

Chapter 1

INTRODUCTION

1.1 Background

Under-five mortality (U5M) is a public health concern worldwide since it is a sensitive indicator of the overall performance of a country. The under-five age group is known to depend on the socio-economic conditions of their environment compared to other age-groups. Due to the high global U5M rate estimated at 93 deaths per 1000 live births in 1990 (1), the Millennium Development Goal (MDG) number 4 was set to reduce U5M by two thirds (67%) between 1990 and 2015 (2). This goal was reviewed at the end of 2015 and was translated to Sustainable Development Goal (SDG) number 3. SDG number 3 aims to reduce neonatal mortality to 12 deaths per 1000 live births and the under-five mortality rate to 25 deaths in 1000 live births by 2030 (3).

The average U5M in Sub-Saharan Africa (SSA) was estimated at 76 in every 1000 live births in 2017, which was 20 times higher than in developed countries' regions with the lowest under-five mortality like Australia and New Zealand and 14 times higher

than in high-income countries in Europe and Northern America (1). Zimbabwe is one of the SSA countries which failed to attain MDG number 4. The U5M was estimated at 75.8 deaths in 1000 live births in 1990 and only decreased to 70.7 deaths in 1000 live births in 2015 (4). This decrease is way too far to reach the set targets.

Most SSA countries failed to meet MDG number 4 because of poverty, corruption, poor allocation and distribution of resources and misplaced priorities (5). Studies have shown that the variation in the occurrence of childhood mortality is associated with household socio-economic factors as well as geographical factors (6). The economic hardships faced by low-income countries greatly affect resource mobilization and shortages of resources have a direct negative impact on the under-five child mortality rates. It is, therefore, crucial to determine some of the factors associated with child survival in Zimbabwe and how these factors changed between 2011 and 2015 and compare the spatial patterns of U5M between 2011 and 2015 as the country targets to achieve SDG 3 by 2030.

Since factors associated with childhood mortality are multifaceted and likely to operate in various complex ways, an appropriate statistical methodology that addresses these complexities are required. In particular, socio-demographic and economic patterns of child mortality vary greatly from place to place and over time. Traditionally, classical statistical models such as correlation coefficients and regression models have been used to model child mortality (7,8) and only produce summary statistics and measures of the association at one particular site but ideally, it cannot be assumed that these relationships hold everywhere within a country. That is, space (spatial) effect of the U5M is not considered in such approaches but it is known that the socio-economic and demographic factors vary from space to space and over time (9). Moreover, one of the limitations of these models is

that the models do not take into account the time contribution of a child. That is; a week-old infant is assumed to have the same contribution as a child who is a few weeks to his/her fifth birthday. It is imperative to note that children do not have the same survival time, some succumb to death very early in life, some experience the event later while some may live past five years. Frail individuals are usually at higher risk of experiencing death earlier than those who are not (10), and frailty may be due to several factors. One of the factors is frailty attributed to unmeasured spatial effects and can be shared by individuals from the same region. This study aimed to account, for spatial effects of U5M rates using a geo-additive Bayesian model, in survival data to assess spatial heterogeneity of U5M and identify possible risk factors of childhood mortality. The proposed model assumes Gaussian Markov Random Field (GMRF) priors for the spatial covariate and estimation is based on the appropriate Markov chain Monte Carlo (MCMC) simulation techniques.

The geo-additive model was fitted to determine the spatial heterogeneity in under-five mortality rates and other factors associated with U5M in Zimbabwe using the Zimbabwe Demographic Health Survey (ZDHS) data collected in 2011 and 2015. The spatial clustering of U5M rates was assessed at the district level between these two-time points. Understanding any spatial clustering of U5M at the district level may help to come up with targeted interventions, guide policy formulation and may give a general guide in the allocation of scarce resources to where they are direly needed especially in low-income countries.

1.2 Problem Statement

Zimbabwe is one of the countries in the SSA region trailing in achieving SDG 3 by the year 2030. U5M rates have been decreasing at a very slow pace and a reduction of

only 6.7% was attained between 1990 and 2015. The recent 2015 ZDHS estimated the U5M rate at 69 per 1000 live births (11); while the U5M rates estimate is decreasing, the childhood deaths are still considered relatively high. The SDG predictions show that Zimbabwe will not be able to achieve the remaining gap by 2030 due to its economic turmoil currently affecting the country (1). Most of the earlier studies conducted on U5M rates have focused on the classical statistical methods which ignore the time components making use of simple regression models (7,8,12,13) and those studies which accounted for the time to examine socio-demographic, economic and health-related risk factors in specific settings of U5M have generally failed to incorporate spatial aspects (9,14,15). However, the U5M burden may vary across regions.

Moreover, such investigations which ignore the spatial aspects of U5M and only focus on the socio-demographic, economic and health-related risk factors usually explain very little of the variation of mortality rate in children. Such modelling approaches may not provide conclusive findings useful for policy formulation as results are generalized to all the country's regions and at times biased estimates may be reported. Therefore, using a geo-additive Bayesian model, in survival data assuming Gaussian Markov Random Field (GMRF) priors for the spatial covariate and estimation based on the appropriate Markov chain Monte Carlo (MCMC) simulation techniques may be a better modelling approach for such data to understand the determinants of under-five mortality in Zimbabwe at the sub-national level.

1.3 Justification

Infant and child mortality has been studied to a considerable extent in SSA and most studies have applied classical approaches using descriptive statistics and regression modelling which do not take into account the proximity of participants from the same

area and also do not consider the time it takes for a child to die. These standard statistical models ignore the spatial dimension in childhood mortality resulting in population-level socioeconomic variables and health resources not explaining fully the mortality rates. However, the inclusion of spatial modelling gives an in-depth overview of how the U5M burden varies from space to space.

Very few studies have been done in Zimbabwe to model child survival; however, to our knowledge none of the studies has made use of geo-additive spatial survival Bayesian models which consider spatial frailty. The geo-additive spatial survival Bayesian models do not only pinpoint the marginalized regions but also hotspots where under-five mortality is high. These models help to identify any regional inequalities across the country to save more children's lives by ending any preventable childhood deaths, which are a global priority.

Moreover, very few studies have studied this area using ZDHS which is a rich database to provide baseline information in a country. Usually, the published ZDHS report only provides information on the aggregated level which masks the spatial variations; hence, it is important to reveal these spatial variations of U5M at a lower administrative level in a region. Findings from this study will be useful in health planning and resource allocation at the sub-national level. The results on the distribution of U5M in different regions will guide policymakers as they come with region-specific interventions to mitigate the problem.

1.4 Conceptual Framework

To identify U5M determinants for this study, we considered the Mosley and Chen framework (Figure 1.1), which is a widely used framework in studies on child survival and was developed by Henry Mosley and Lincolin Chen in 1984 (16). Mosley and Chen suggested that U5M was a result of interaction between biological factors and socio-economic factors over time and hence U5M should be considered as a chronic process rather than a sudden event. (16). In the framework, the biological factors were treated as the proximate factors and were divided into maternal factors (maternal age at childbirth, contraceptive use, maternal BMI, birth interval, parity); environmental (pollution); nutrition; injury; and illness control (prohibition and therapy). The Socio-economic factors were treated as distant factors and handled at the individual level (culture, attitude, religion), household level (wealth status) and community level (service delivery, economy) (16,17).

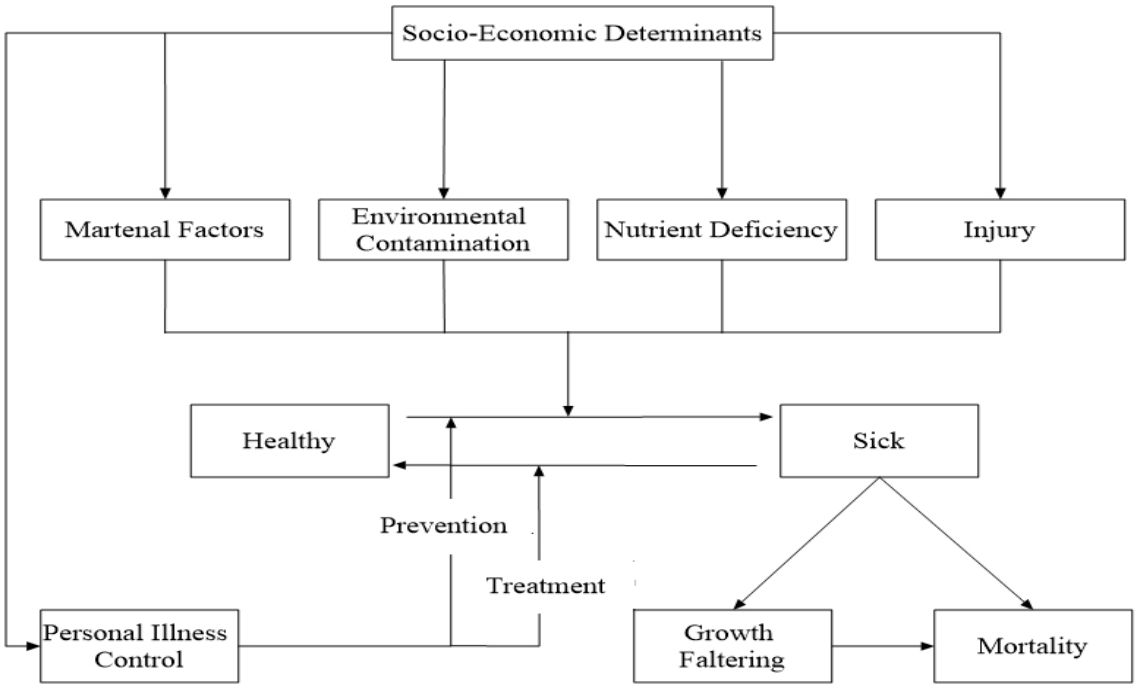


Figure 1.1: Mosley and Chen framework of child survival

The conceptual framework for the current study (Figure 1.2), which links U5M to

various derteminants was adapted from the conceptual framework developed based on the Mosley-Chen framework by Lamichhane et al. in 2017 (17). The factors associated with U5M were grouped into community level, socio-economic and proximate determinants (maternal and child-related factors). The proximate factors had a direct effect on U5M while socio-economic and community factors indirectly affected U5M.

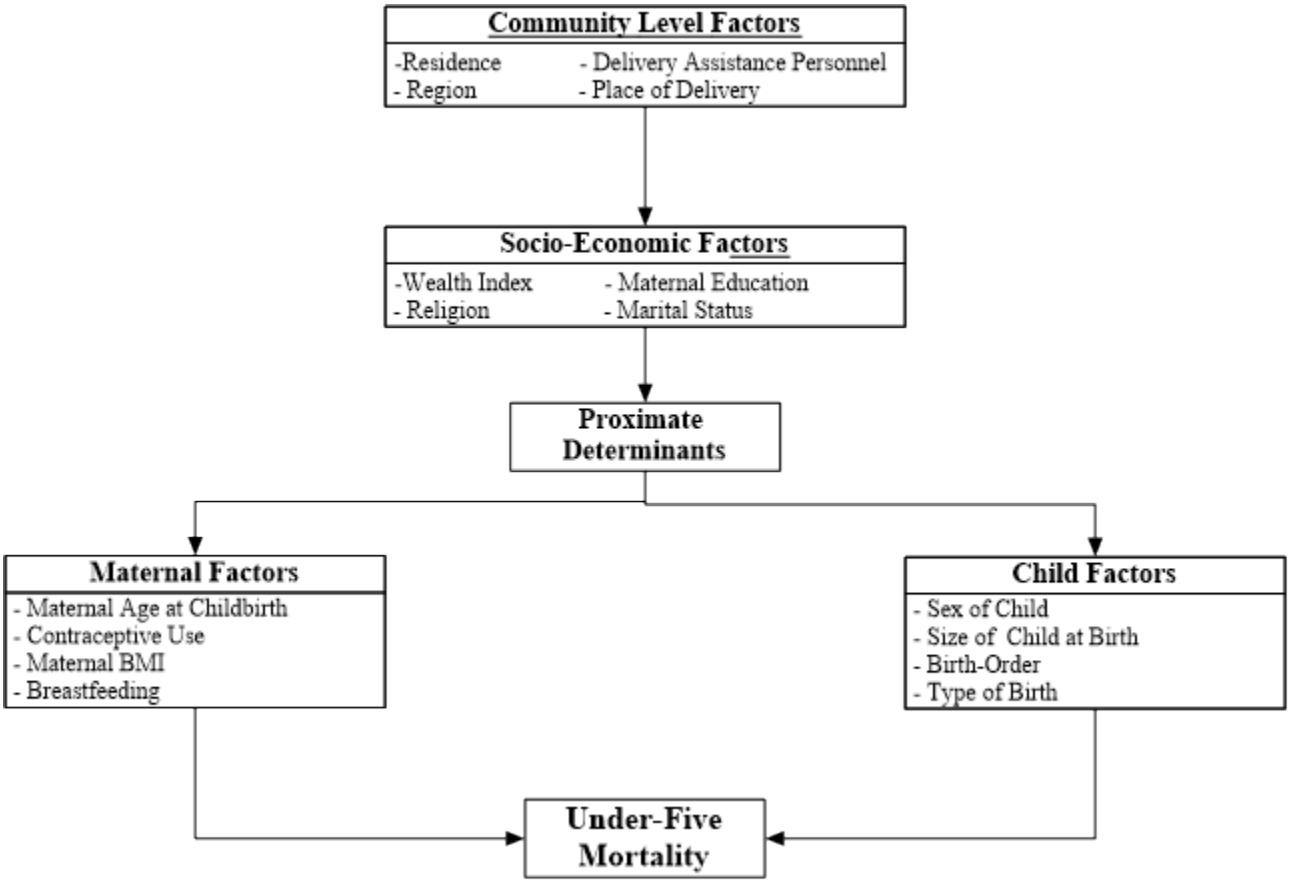


Figure 1.2: Conceptual framework of under-five mortality

1.5 Literature Review

The burden of U5M remains a global public health concern; however, substantial efforts have been made by governments and other organisations such as the United Nations (UN) to reduce the U5M rates over the past decades. Despite this positive global progress, uneven progress has been observed in developing countries which

are still lagging due to various socio-economic challenges faced by these countries (18). Though some developing countries have shown a decrease in U5M rates in the past years, the decline has been of minimum to negligible magnitude in most countries (18). SSA still accounts for most of the world's U5M with an average of 76 deaths per 1000 live births (1). The U5M has disproportionately affected SSA countries as U5M rates are lower in Southern parts of the region and increases as you drift away from Southern Africa towards West and Central Africa (18).

Extensive research has been done on infant and child mortality and associated risk factors in SSA and several factors are associated with U5M. For child-specific factors, baby size at birth proved to be a significant risk factor of U5M and small babies (birthweight < 2500g), were more likely to die than normal-sized babies (birthweight ≥ 2500g)(19–21). A condition called fetal macrosomia which is defined as large baby size at birth (birthweight ≥ 4000g) and is caused by factors like high maternal age, genotype, diabetes and high maternal BMI has also been shown to be one of the driving factors of U5M in developed countries such as Turkey and the USA (22,23) and this might also be the case in SSA.

Nutrition was critical in reducing the risk as those infants and babies who are malnourished had an increased risk of death (24). Azuine et.al investigated the association between exclusive breastfeeding and U5M and suggested that the nutritional benefits of breastfeeding were essential for the full development of a child and the antibodies in breast milk provided immunity against infections and diseases which was protective against mortality (25). As far as the sex of a child is concerned, U5M has been reported to be significantly higher in boys compared to girls and the explanation for this might be that biologically boys had a weaker immune system than girls (26,27). From literature, multiple births were a risk factor of child mortality

when compared to singletons (28) and findings from a study in Bangladesh suggested that multiple births were associated with complications at birth, higher risk of having defects and disabilities and being small in size all of which decreased the chances of survival (29).

Previous studies have investigated the effect of maternal factors which include parity, birth spacing, contraceptive use, antenatal care, postnatal care, maternal age on infant and child mortality. As we would anticipate, there is a very close relationship between U5M mortality and maternal factors; however, the effects of these factors decrease as the baby grows older (15,27). Kembo and Van-Ginneken stressed the importance of maternal age in child survival and reported that young mothers (less than 20 years) in Zimbabwe had a higher chance of losing their child in infancy as they were more prone to experiencing delivery complications and also lacked the necessary knowledge and experience on raising a child (15). Contradictory results have been obtained on the effect of maternal education on infant and child mortality since some studies have reported a strong association between education and U5M (8,12) while other studies have reported an insignificant relationship between maternal education and U5M (30).

A study conducted in Zimbabwe reported that higher education attainment of the mother improved child (those aged one year and above) survival but had little or no effect on infant (those aged less than one year) survival (15). Antenatal care and postnatal care are very crucial factors in reducing the risk of U5M as reported in some studies (13). The more the clinical visits of the mother, the lower the chances of infant mortality. Family planning or contraceptive use by women is usually done by people who are concerned about their wellbeing and plan their families.

In a study done in Kenya and Zimbabwe on the effect of contraceptive use on U5M, it was shown that contraceptive use improved preparedness, reduced unplanned pregnancies, lowered chances of high-risk births, increased birth interval; hence lowering U5M risks (31). In an investigation on the effect of socio-economic factors on U5M, Kamniki et al. reported that married mothers were less likely to experience a loss of a child. Reasons being that married women got financial and social support from their partners and marriage allowed for bringing together resources to provide for the child's wellbeing (7).

It is evident from the literature that socio-economic activities are associated with child survival and the use of cleaner fuels, improved sanitation (piped water and toilet facilities) and high wealth status have been proven to significantly reduce infant and child mortality in SSA (8,15). In his paper on childhood diseases in Zimbabwe, Tsiko suggested that the disparities which exist in U5M between nations extend to the sub-national level, and variations are observed in regions within the same country as provinces and districts possess diverse environmental conditions such as climate (rainfall, temperature, wind), soils which are associated with infant and child morbidity and mortality (26). Furthermore, economic activities such as agriculture, mining, etc. also affect the environment and these are practised at different magnitudes across the country, and they have played a significant role in childhood morbidity and mortality (26,32).

In terms of location, rural areas and urban areas might share similar geographical locations but the environmental conditions are different. A child born in an urban area has been shown to have a higher chance of survival compared to a rural counterpart and also a child born in a conflict-free environment has a better chance of survival compared to a counterpart in a conflict infested region (6). Religious affiliations and cultural beliefs play a crucial role in U5M as age at sexual debut, age at marriage and

the peoples' willingness to use healthcare facilities depend on culture and religion and these factors are strongly associated with child survival (33).

1.5.1 Frailty in Survival Analysis

The above risk factors are due to measured factors of under-five mortality but some factors cannot be easily measured and accounting for such factors in the study may give better results, since unmeasured effects introduce some frailty in the populations. Infants and children who are frail are at a higher risk of dying before their fifth birthday. In this context, frailty may be defined as some state of vulnerability attributed to unmeasured effects such as genetic conformation, lifestyle, environmental factors, etc. which cannot be distinctly measured or uniquely identified in individuals (34). In consideration of frailty survival analysis, we are interested in the different survival experiences of individuals and the variability in time to the event which emanates mainly from two sources which are fixed effects and from the frailty random effects (34). Frailty can be at the individual level; the univariate case or can be common amongst groups of people or clusters in which case it is called the multivariate case (35). Naturally, we would expect individuals from neighbouring regions to experience similar survival times due to similar environmental and service delivery factors. As a result, in statistical inference, we should consider fixed effects and the random effects (frailty) due to spatial variations as there is a high correlation in survival times of individuals from the same region because of shared frailty (36).

1.5.2 Bayesian Spatial Survival

Frailty models have been important in survival analysis and have been used to fit models on infant and child survival where most authors preferred Cox-PH models with frailty (37–39). Survival analysis using Cox-PH has been used to model U5M in Zimbabwe and examples are Kembo and Van-Ginneken who used Cox-PH to model

infant and child mortality using 2005-06 ZDHS data (15) and Yaya et al. who did a similar study but used 2015 ZDHS data to model U5M and associated maternal factors (19). In both studies, maximum likelihood estimation (MLE) was used and the geographical variations of U5M were not considered.

In SSA, a few studies on spatial survival of under-five children have been done; Gayawan et al. applied a fully Bayesian approach and used geo-additive discrete-time survival modelling to determine spatial variations of U5M in West Africa using 2008-2013 DHS data. (32). Kandala and Ghilagaber also applied geo-additive Bayesian discrete-time spatial survival modelling to study geographical differences of U5M at the district level in Malawi using 2000 DHS data. (40). Cox-PH with frailty model was used by Kazembe et al. to model cross-sectional census data on geographical inequalities of infant and child mortality in three SSA countries (Rwanda, Uganda and Senegal) (41). Similarly, a study conducted in Nigeria used Cox-PH to model determinants associated with U5M using 2013 DHS data (9).

Survival analysis with the spatial random effects has not been applied much to U5M in SSA. In this study we were interested in using survival analysis models to incorporate frailty random effects and determine the geographical variations of U5M after adjusting for associated factors. The use of cross-sectional DHS data from a one-time point in survival analysis was demonstrated in the studies above and, in this study, we sought to use ZDHS data in the same way but from the last two surveys (2011 and 2015), to monitor the progress of child survival in Zimbabwe. To achieve this, we explored the use of OpenBugs Bayesian statistics software and R statistical software to fit the models. OpenBugs is well adapted for time-to-event spatial data and has over the past been used to fit Cox-PH (42,43) and AFT (44–46) models. More recently, a package has been developed to fit Bayesian spatial survival models (Cox-PH, AFT and PO) in

R statistical software (47) and it is worth using in modelling U5M.

1.6 Research Question

What are the determinants and geographical variations of under-five mortality in Zimbabwe between 2011 and 2015 survey points?

1.7 Objectives

1. To describe and compare under-five mortality rates by background characteristics between 2011 and 2015 survey points.

2. To determine and compare associated risk factors of under-five mortality in Zimbabwe between 2011 and 2015 surveys points.

3. To determine and compare the spatial heterogeneity (structured and unstructured effects) of under-five mortality adjusting for possible risk factors in Zimbabwe at the district level between 2011 and 2015 surveys points.

Chapter 2

METHODS

2.1 Introduction

This chapter outlines the methods which were used for data collection, extraction, transformation and analysis. It starts by describing the study design, study site and population. It then outlines the data management steps including data extraction, cleaning and recoding and lastly the spatial survival analysis techniques which were used to model U5M are described.

2.2 Study Design/Site

Secondary data analysis based on the ZDHS under-five children's data collected in 2011 and 2015. Hence, this study was conducted in Zimbabwe. Zimbabwe is an inland country bordering South Africa, Botswana, Zambia and Mozambique. The country is partitioned into eight provinces (Manicaland, Midlands, Masvingo, Matabeleland South, Matabeleland North, Mashonaland East, Mashonaland West and Mashonaland

Central) and two metropolitan cities (Harare and Bulawayo) with Harare as the Capital City of the country (Figure 2.1). These provinces are further portioned into districts giving a total of 60 districts (Figure 2.2) which further break down into more than 1500 wards.

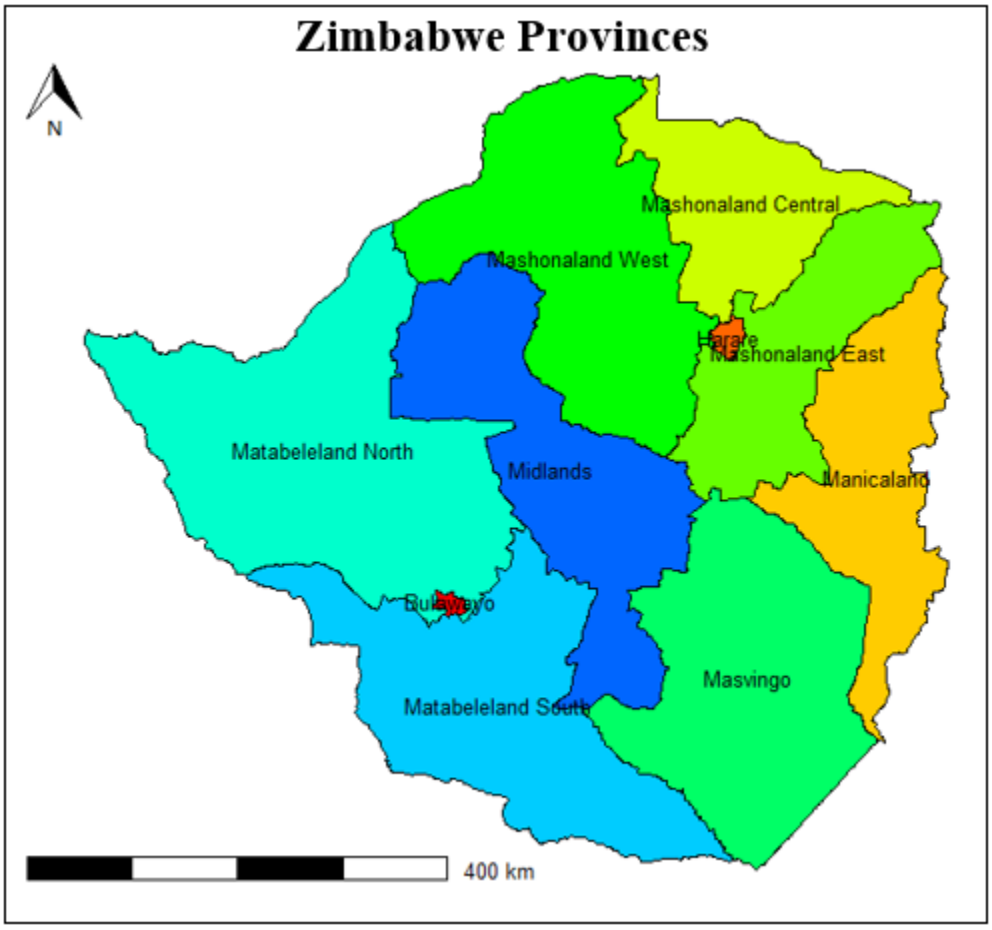


Figure 2.1: Map showing Zimbabwe provinces labelled by names

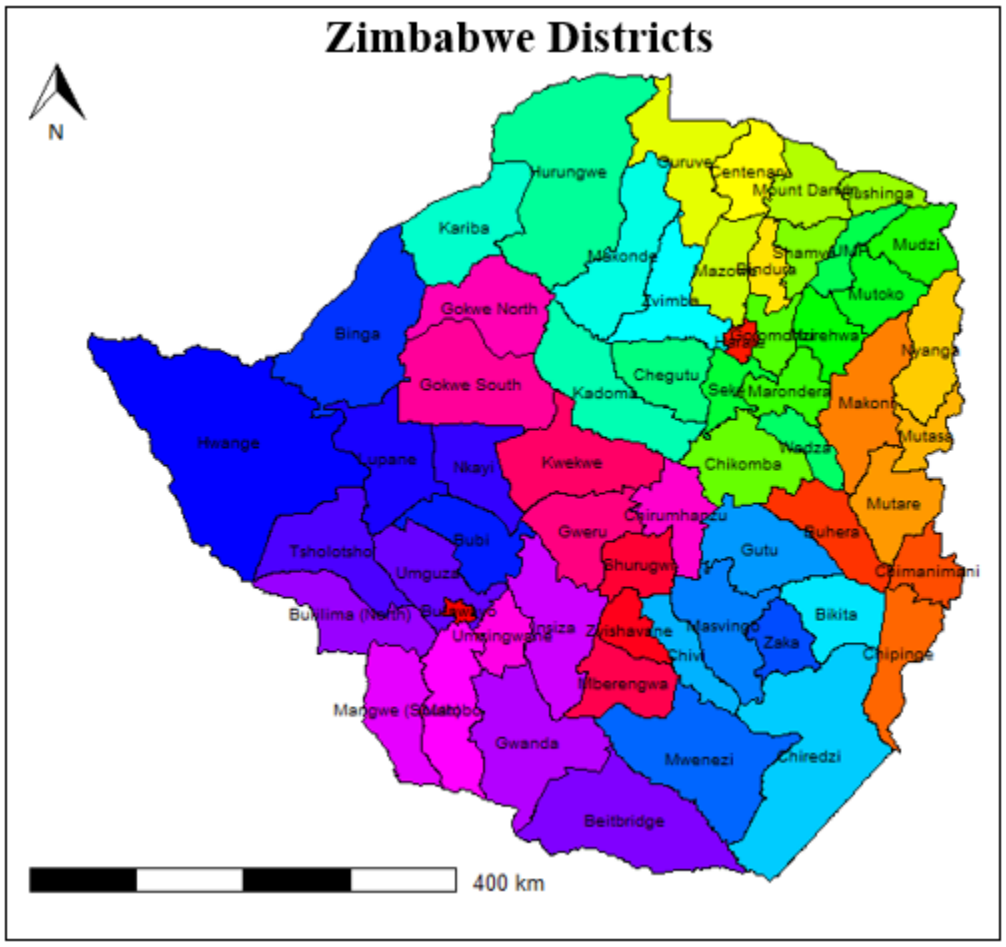


Figure 2.2: Map showing Zimbabwe districts labelled by names

2.3 Data Sources

The data used for this study came from the Zimbabwe national Demographic Health Surveys conducted in 2011 and 2015. The surveys are done countrywide after every five years to obtain the latest quality data which has been collected using expert techniques to monitor demographic and health indicators. The data was collected through individually administered questionnaires. The questionnaires collect information on maternal and child health, reproductive health and fertility, HIV/AIDS, immunization, nutrition, child mortality and household data. In the women's questionnaire, some of the information collected is on birth history and

childhood mortality which this study focused on (11,48).

The sampling technique used in these surveys was the multistage stratified sampling. The two-staged stratified sampling process is done to identify the particular household to be considered in the survey. Firstly, the stage was the selection of enumeration areas (EA) which are embedded within wards (11,48) . In 2011, 406 EAs were selected of which 169 were in urban and 237 in rural areas (Figure 2.3). In 2015, 400 EAs were selected of which 166 were in urban and 234 in rural areas (Figure 2.3). The second stage involved the selection of households; in 2011, a total of 9,756 households were interviewed while in 2015, 10,535 households were interviewed (11,48). From these selected households, 9171 eligible women (15-49 years) and 7480 eligible men (15-49 years) were successfully interviewed in 2011 (48). Similarly, for the 2015 survey, a total of 9955 eligible women and 8396 eligible men were successfully interviewed (11). The survey also collects information for children aged 0-59 months from women of the reproductive age group. The detailed information of sampling instruments, sampling techniques, data collection and management used by ZDHS is publicly available (https://dhsprogram.com/).

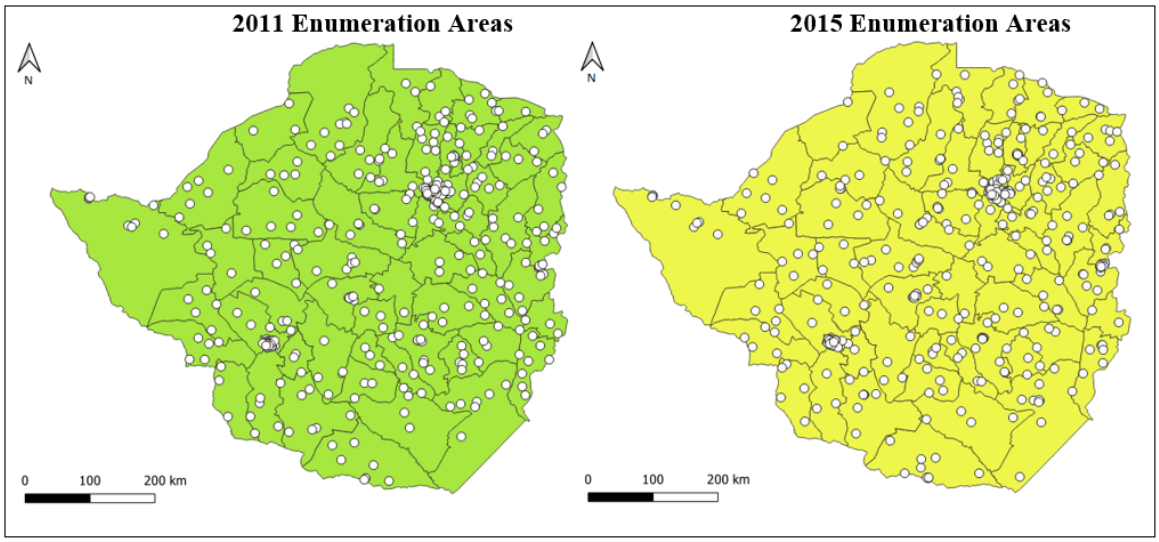


Figure 2.3: Enumeration (cluster) points for 2011 (left) and 2015(right) ZDHS.

2.4 Data Quality

The DHS datasets come from a standardized survey done in many countries every five years. The datasets are cleaned and compiled after every survey and stored in different files for each country namely household, household members, women, men, HIV, births, couples, and children including the under-five children's data files that were used in this study. The data collected is of high quality, as interviews are conducted by specially trained personnel on participants with high response rates across all the country's regions. Due to the design of the survey, the variable information is weighted to represent the true population from which the data is collected at the cluster level. Standard and consistent datasets are stored and comparisons between two survey time-points can easily be done (49). Although the data is of high quality, it is not perfect and some of the downfalls of DHS data are that different EAs are used every time a survey is done, and this might affect comparability. There are also sensitive questions asked in the surveys and these may not always be answered correctly.

2.5 Current Study Population

This study targeted under-five children aged 0-59 months both alive and dead as recorded in the 2011 and 2015 survey data. Only live births were considered, and any record of still birth and miscarriages excluded. The final sample sizes for the under-five children were 5,563 and 6,132 in 2011 and 2015, respectively.

2.6 Data Management

2.6.1 Data Extraction

Outcome Variable

The U5M information of children born 0 to 59 months before the survey was extracted from the birth history section of the individual women file. For each live birth to a woman, information on year and date of birth, whether the child is dead or alive, the age at death (if the child was dead), sex of the child was captured. The U5M rate was estimated from the data in the birth history from a child's birth, survivorship status, and age at death (if dead). The outcome variable for this study was the risk of under-five death which was measured from the duration of survival in days. This was defined as the risk of a child dying before reaching its fifth birthday. The survival status of infants was further recoded as '1' for the infant who died within the first 59 months of life and '0' for infants who survived beyond 59 months of life.

Independent Variables

The variables used in this study only came from the children's data file. Literature guided variable selection and the variables extracted include cluster number, sample weight, date of death, date of birth of the child, date of the interview, mother's age at birth in years (continuous variable), child's birth-weight in grams (continuous variable), and size (small, average and large at birth), contraceptive use (yes or no), religion (Christianity, Pentecostalism, Apostolic and the rest classified as other), wealth index (poor, average and rich) , region, mother's marital status (single, married and widowed/divorced), birth type (singleton or twin (multiple)), mother's education(no education, primary and secondary/higher level), sex of child (male or female) and type of residence (rural or urban).

2.6.2 Data Preparation

The extracted data was cleaned by checking for completeness of the information and checking how the variables were coded. Entries with numerical codes for missing values such as child's birth weights were re-coded to missing. Recoding of categorical variables and renaming of variables was done for ease of data handling. New variables were generated from some of the existing variables. For instance, the wealth quantile variable had five categories which were further reduced to three categories for easy data interpretation. The outcome variable named "child alive" in the ZDHS was reverse coded such that "0" indicated a child who was alive and "1" indicated a child who died. Time to death in days was calculated using the date of birth and date of death if the child died while those who were alive were censored based on the date of interview from the date of birth.

2.6.3 Missing Data

Exploratory data analysis was used to determine the data structure for each variable and identify any missing values. Most of the variables we selected were complete and from the 2011 dataset, contraceptive use and size of the baby at birth were the two variables with missing data while from the 2015 dataset, the size of the baby at birth had missing values. The data were assumed to be missing at random (MAR) and multivariate imputation by chained equations (MICE) package in R statistical software was used for imputation of the missing values (50). A sensitivity analysis (Appendix B) using the standard Cox proportional hazard model on both complete case data and imputed data showed that imputed data did not improve the model estimates as the missing percentages were below 3%; hence, the imputed data were very similar to complete case data and inferences were based on this data.

2.7 The MICE method

Missing data imputation was done using the *mice package* in R statistical software, developed by Stef van Buuren and Karim Groothuis-Oudshoom in 2011 (50) and it uses a Gibbs sampling algorithm in MCMC technique. MICE method needs a few iterations and is done variable by variable on target columns as the complete variables are used to predict the missing values. From the selected data, two variables had missing values from the 2011 dataset, and these were the size of the child at birth and contraceptive use while the size of the child at birth was the only variable with missing values from the 2015 dataset. There was a total of 150 missing values in the 2011 dataset of which 12/5563 (0.22%) were for contraceptive use and 138/5563 (2.75%) were for size while there were 26/6132 (0.42%) missing values for size at birth in 2015 dataset.

Imputation of missing values was run using five imputation models, and to achieve convergence a maximum of 50 iterations were done under the assumption that data were MAR. The size of the baby at birth was a categorical variable with three ordered classes (large, average and small); hence, the proportional odds model for ordinal data was used as the imputation method. In contrast, the contraceptive variable was imputed using the logistic regression imputation method since the variable was binary.

2.8 Spatial Data

The polygon shapefile data for the Zimbabwean map was downloaded from the DIVA shapefile site (https://www.diva-gis.org/) (51). A boundary shapefile with 60 administrative levels (districts) was used to define the spatial unit of analysis in this study. The Zimbabwe centroid data shapefiles with the specific EA coordinates corresponding to the survey cluster variable were obtained from the DHS website (https://dhsprogram.com/) (52).An open-source system, Quantum Geographic In-

formation Systems (QGIS) (53) was used to match the clusters (centroid shapefile) and spatial polygon shapefile (Figure 2.3 and 2.4). For spatial analysis in R, the shapefiles were loaded directly as R can read shapefiles (shp format) but for spatial analysis in OpenBugs, we converted the shapefiles in R to S-plus format and exported them to OpenBugs.

2.9 Statistical Analysis

Objective 1: *To describe and compare under-five mortality rates by background characteristics between 2011 and 2015 survey points.*

Descriptive statistics, that is, frequencies and percentages for categorical variables were used and continuous variables were summarised by means (SD) as the data were normally distributed using STATA version 15.0. Univariate analysis of survival experiences by the group was visualised using non-parametric Kaplan-Meier (KM) plots and analysis was done in R statistical software version 3.6.2 (54) using the *survival* package for analysis of survival times and *ggplot2* and *fortify* packages for plotting the survival experiences with time (55). The log-rank test was used to determine the difference in survival rates between an independent group of variables.

Objective 2: *To determine and compare associated risk factors related to under-five mortality in Zimbabwe between 2011 and 2015 survey point.*

Semi-parametric and parametric geo-additive Bayesian survival modelling approaches were used for the study. Semi parametric Cox proportional hazards (Cox-PH) model was used with and without the spatial term. The MLE approach was used for the non-spatial Cox-PH model while for the spatial models, the geo-additive Bayesian survival model with frailty terms was used with prior information assigned to all unknown parameters. Similarly, parametric AFT was fit with and without a spatial

term. The parametric geo-additive Bayesian models assume the baseline hazard function to be either log-logistic, Weibull or log-normal accelerated failure time (AFT) models with frailty. The models were fitted in OpenBugs Bayesian software version 3.2.3 (56) and R statistical software. The motivation behind using the Cox-PH and AFT models was based on that the Cox-PH gives the chances of a child dying as hazard ratios and the AFT models give the differences in terms of the number of days survived by a child in the study and AFT can be used in situations where the proportional hazards (PH) assumption does not hold hence, use of these approaches results in more informed conclusions. The Cox-PH and the AFT geo-additive models were modelled based on the Bayesian estimation perspective. The model specifications are described in detail in the next section. In this study, Cox-PH and AFT models were fitted in R and OpenBugs. However, since their specifications are slightly different, in the next section, the model specifications for OpenBugs are explained while the model specification for the R models are explained in Appendix C.

Objective 3: *To determine and compare the spatial distribution of under-five mortality risk factors in Zimbabwe at the district level between 2011 and 2015 survey point.* Geo-additive Bayesian survival models with frailty terms, were fit with prior information assigned to all unknown parameters. The shared frailties captured the spatial component (structured and unstructured spatial effects) of the models. The models were adjusted for spatial effects and U5M associated factors and fit in R and OpenBugs software. The corresponding maps for structured and unstructured spatial effects for the two time points were then plotted in both software.

2.10 Kaplan-Meier Model

The Kaplan-Meier method is a non-parametric method that was developed by Edward L Kaplan and Paul Meier to analyse time-to-event data of which in our case it is the time to death of a child (57,58). KM is best used for univariate analysis of survival experiences for categorical variables and plotting the survival curves which show steps going down for every event experienced giving better visualisation of the different survival times between groups and the significance of the differences maybe be confirmed using the log-rank test (57,59). Since KM is a univariate method, we can neither use it for continuous variables nor adjust for the effect of other variables and potential confounders on the survival times but none the less it can be used for preliminary analysis which gives one an idea of survival experiences of the participants (57).

2.11 Cox Proportional Hazards Model Specification

The Cox-PH model is analogous to the KM model, but it allows the fitting of more than one variable including continuous covariates. The method was presented by Sir David Cox in 1972 (60) and has been over the years used for time to event studies and its flexible nature has allowed researchers to fit semi-parametric and parametric Cox-PH models. The coefficients of Cox regression are exponentiated and the results given as the hazard ratios under the assumption that the hazards between groups remain constant over time and this is called the proportional hazards assumption. In our context, the hazard may be defined as the probability that a child dies the next instance given that the child survived till a certain time point, (t), that is hazard

$(h(t))$ maybe written as:

$$h(t) = lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}, \tag{2.1}$$

where (T) is time to the event (time to death of an under-five child) and $t + \Delta t$ defines the next instance such that Δt is a small increment in time. Equation 2.1 can be written as:

$$h(t)dt = P(t \leq T < t + \Delta t \mid T \geq t) \tag{2.2}$$

$$= \frac{P(t \leq T < t + \Delta t \cap T \geq t)}{P(T \geq t)} \tag{2.3}$$

$$= \frac{P(t \leq T < t + dt)}{P(T \geq t)}. \tag{2.4}$$

If we want to determine the effect of covariates on the survival time, we use the following hazard function.

$$h(t \mid X) = h_0(t) exp(\boldsymbol{\beta}\mathbf{X}), \tag{2.5}$$

where $h_0(t)$ is the overall baseline hazard, that is, the hazard when the effect of covariates is null, $\boldsymbol{\beta}$ is a vector of regression coefficients and $\mathbf{X}$ represents the matrix of the covariates. Equation 2.5 maybe extended by including spatial frailty shared by children from the same district (w):

$$h(t) = h_0(t) exp(\beta\mathbf{X} + w). \tag{2.6}$$

In modelling, we may take logs of both sides for additive effect and remain with the following expression:

$$log(h(t \mid X)) = log(h_0(t)) + \boldsymbol{\beta}\mathbf{X} + w. \tag{2.7}$$

Which can be simplified to:

$$\lambda(t) = \lambda_0(t) + \boldsymbol{\beta}\mathbf{X} + w\,, \tag{2.8}$$

where $\lambda(t)$ is the rate of an event occurring (rate of U5M). In semi parametric models the survival times and the baseline hazard are assumed not to follow any distribution, but the effect of the covariates is assumed to follow some distribution.

Likehood Function

The following likelihood can be used for a right-censored time to event data. Let survival time for child i be T_i, the density function be $f(T_i)$, the survival function be $S(T_i)$ and δ_i be the survival status for child i such that:

$$\delta_i = \begin{cases} 1 & \text{if child } i \text{ is dead} \\ 0 & \text{if child } i \text{ is alive}\,. \end{cases}$$

The likelihood of survival for the Cox PH model will be given by:

$$L(.) = \Pi_{i=1}^{n} \left[f(T_i)\right]^{\delta_i} \left[S(T_i)\right]^{1-\delta_i} \tag{2.9}$$

$$= \Pi_{i=1}^{n} \left[h(T_i)S(T_i)\right]^{\delta_i} \left[S(T_i)\right]^{1-\delta_i} \tag{2.10}$$

$$= \Pi_{i=1}^{n} \left[h(T_i)\right]^{\delta_i} \left[S(T_i)\right]\,. \tag{2.11}$$

We adjusted the model used in OpenBugs and adapted our data accordingly. The code is developed using the counting process and is best described in the OpenBugs manual (56) and well implemented in (42,43) hence we will follow the steps as we describe the model.

Let $i = 1,...,n$ be the number of subjects, in our case children, then $N_i(t)$ is the count of the number of deaths up to some time t. We may then write the hazard rate or

intensity of failure ($\lambda_i(t)$) as:

$$\lambda_i(t) \;=\; E(dN_i(t) \mid F_{t-})\,, \tag{2.12}$$

where $dN_i(t)$ is a small increment of N_i in the interval $[t, t+\Delta t]$.The data/history available before time t is F_{t-} and since $dN_i(t) = 1$ if subject i dies at time t and $dN_i(t) = 0$ if subject i survives past t, hence $E(dN_i(t) \mid F_{t-})$ is the probability of death in the next instance for child i.

The hazards may then be written as:

$$\lambda_i(t) \;=\; Y_i(t)\lambda_0(t)exp(\beta^T\mathbf{X}_i + u_j + v_j)\,, \tag{2.13}$$

where $Y_i(t)$ indicates the survival status at time t and takes the value 1 if the child is still alive and 0 otherwise, $\mathbf{X}_i$ are the explanatory variables for child i and u_j represents the unstructured spatial random effects while v_j represents the structured spatial random effects or the shared frailty for $j = 1,...,m$ districts.

Now let $\Lambda_0 = \int_0^t \lambda_0(u)du$ be the cumulative baseline hazard and let the data, $D = N_i(t), Y_i(t), \mathbf{X}_i$ for $i = 1,...,n$.

The likelihood can then be written as:

$$L(D \mid \beta, \Lambda_0, v) \;=\; \Pi_{i=1}^n \left[h(T_i)\right]^{\delta_i} \left[S(T_i)\right] \tag{2.14}$$

$$\;=\; \Pi_{i=1}^n \left[\Pi_{t\geq 0}\lambda_i(t)\right]^{dN_i(t)} exp(-\lambda_i(t)dt) \tag{2.15}$$

$$\;=\; \Pi_{i=1}^n \left[\Pi_{t\geq 0}\lambda_i(t)^{dN_i(t)}\right] exp(-\lambda_i(t)dt)\,. \tag{2.16}$$

Prior Specification

Since $N_i(t)$ is the count of dead children at time t, it means that $N_i(t) \sim Pois(\mu)$ hence $dN_i(t)$ will be the count of the small increments in the intervals $[t, t+\delta t]$ with

mean $\lambda_i(t)dt$.

$$dN_i(t) \sim Pois(\lambda_i(t)dt)\,, \tag{2.17}$$

where:

$$\lambda_i(t)dt = Y_i(t)exp(\boldsymbol{\beta}^T\mathbf{X}_i + v_j)d\Lambda_0(t), \tag{2.18}$$

and $d\Lambda_0(t) = \Lambda_0(t)dt$ is a small increment on the baseline hazard in the interval $[t, \Delta t]$. The gamma prior for the cumulative baseline hazard will fit in well since it is a conjugate prior to the Poisson distribution.

$$d\Lambda_0(t) \sim Gamma(cd\Lambda_0^*(t), c)\,, \tag{2.19}$$

where $d\Lambda_0^*(t)$ is a prior guess of the hazard which is probably unknown, and c is the degrees of freedom of this prior guess.

$$\Rightarrow p(d\Lambda_0) = \frac{c^{cd\Lambda_0^*}}{\Gamma(c^{d\Lambda_0^*})}\Lambda_0^{(cd\Lambda_0^*-1)}e^{-c\Lambda_0} \tag{2.20}$$

$$\propto \Lambda_0^{(cd\Lambda_0^*-1)}e^{-c\Lambda_0}\,. \tag{2.21}$$

The prior distribution of the regression coefficients (β) may be assigned a normal distribution with mean μ and variance σ^2; $\boldsymbol{\beta} \sim N(\mu_\beta, \sigma_\beta^2)$.

$$\Rightarrow p(\beta) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}}\exp\left[\frac{-1}{2}\left(\frac{\beta-\mu_\beta}{\sigma_\beta}\right)^2\right] \tag{2.22}$$

$$\propto \exp\left[\frac{-1}{2}\left(\frac{\beta-\mu_\beta}{\sigma_\beta}\right)^2\right]. \tag{2.23}$$

The unstructured spatial random effects u are assigned a Gaussian distribution with

mean = 0 and precision τ_u; $u \sim N(0.0, \tau_u)$

$$\Rightarrow p(u \mid \tau_u) = \frac{1}{\sqrt{2\pi\frac{1}{\tau_u}}} \exp\left[\frac{-1}{2}\left(\frac{u}{\tau_u^{-1}}\right)^2\right] \tag{2.24}$$

$$\propto \exp\left[\frac{-1}{2}u^2\tau_u\right]. \tag{2.25}$$

The structured spatial random effects (v) are assigned an intrinsic conditional auto regressive prior (ICAR). Intrinsic conditional autoregressive models are part of conditional autoregressive models (CAR) presented by Besag and Kooperberg in 1995 (61). These models cannot be used as the likelihood for data and can only be used as prior distribution for spatial effects in Bayesian spatial models (62). ICAR models define a Gaussian Markov random field (GMRF) and assume complete spatial correlation of neighbouring regions (62,63). The main advantage of ICAR models is in the simplicity of specification of parameters and as a result, the same precision (τ) can be used for all regions (63,64). The correlated (structured) spatial effects for all our models were therefore assigned the ICAR prior distribution due to its advantages.

The derivation of the ICAR model from the CAR model is shown below and the steps followed are mainly based on derivation by Morris et.al (62,64) and work by Lavine and Hodges (63).

Let W be an $n \times n$ adjacency matrix for n regions and the spatial effects for the regions given by: $v = (v_1, \ldots, v_n)^T$. Let $i \sim j$ denote that i and j are neighbouring regions. Then v is normally distributed with mean zero and covariance Σ, that is: $v \sim N(0, \Sigma)$ Since $\Sigma = Q^{-1}$, where Q is a positive definite symmetric precision matrix, this is the same as writing $v \sim N(0, Q^{-1})$

The precision matrix is related to the spatial weights and number of neighbours each

region has by:

$$Q \;=\; [\tau(M-W)]\,, \tag{2.26}$$

where τ is the precision, M is a diagonal matrix of number of neighbours for each region such that $m_{i,i}$ gives the number of neighbours for region i and W is the adjacency matrix with spatial weights such that $w_{i,j}=1$ if $i\sim j$ (i and j are neighbours) and $w_{i,j}=0$ otherwise;

$$\Rightarrow v \sim N(0,[\tau(M-W)]^{-1})\,. \tag{2.27}$$

Hence:

$$p(v_i \mid v_j,\tau_i) \sim N\left(\frac{\sum_{i\sim j} v_i}{m_{i,i}\tau_i}, \frac{1}{m_{i,i}}\right) \text{ and } p(v_j \mid v_i,\tau_j) \sim N\left(\frac{\sum_{i\sim j} v_j}{m_{j,j}\tau_j}, \frac{1}{m_{j,j}}\right).$$

The joint spatial distribution is given by:

$$p(v) \;=\; \frac{1}{\sqrt{2\pi[\tau(M-W)]^{-1}}} \exp\left[\frac{-1}{2[\tau(M-W)]}(v^T v)\right] \tag{2.28}$$

$$\propto\; \exp\left[\frac{-1}{2}[\tau(M-W)](v^T v)\right]. \tag{2.29}$$

For neighbouring regions i and j

$$p(v) \;=\; \exp\left[\frac{-\tau}{2}\sum_{i,j}(M-W)v_i v_j\right] \tag{2.30}$$

$$=\; \exp\left[\frac{-\tau}{2}\sum_{i} v_i v_j M_{i,j} - \sum_{i,j} W_{i,j} v_i v_j\right]. \tag{2.31}$$

Since M is a diagonal matrix with only $M_{i,i}$ entries and $w_{i,j} = w_{j,i}$, we get the following

$$\begin{aligned} p(v \mid \tau) &= \exp\left[\frac{-\tau}{2}\sum_i v_i^2 M_{i,i} - \sum_{i\sim j} 2v_i v_j\right] && (2.32)\\ &= \exp\left[\frac{-\tau}{2}\sum_{i\sim j}(v_i^2 + v_j^2) - \sum_{i\sim j} 2v_i v_j\right] && (2.33)\\ &= \exp\left[\frac{-\tau}{2}\sum_{i\sim j}(v_i^2 - 2v_i v_j + v_j^2\right] && (2.34)\\ \Rightarrow p(v \mid \tau) &\propto \exp\left[\frac{-\tau}{2}\sum_{i\sim j}(v_i - v_j)^2\right]. && (2.35) \end{aligned}$$

This is known as the pairwise difference formula and depends on the constraint $\sum_i v_i = 0$ for it to be identifiable (62,63). The ICAR prior is improper and needs the specification of the prior precision τ which follows a gamma distribution; $\tau \sim gamma(\alpha, \beta)$ (62).

$$\Rightarrow p(\tau) \propto \tau^{\alpha-1} e^{-\beta\tau}. \quad (2.36)$$

Hence

$$p(v \mid \tau)p(\tau) \propto \exp\left[\frac{-\tau}{2}\sum_{i\sim j}(v_i - v_j)^2\right]\tau^{\alpha-1}e^{-\beta\tau}. \quad (2.37)$$

The Hyperpriors

The hyper-prior for the unstructured spatial effects was assigned a Gamma distribution $\tau_u \sim Gamma(a_u,b_u)$ with pdf:

$$p(\tau_u) \propto \tau_u^{a_u-1} e^{-b_u\tau_u}. \quad (2.38)$$

The hyper-prior for the structured spatial effects follow a Gamma distribution $\tau_v \sim Gamma(a_v,b_v)$ with pdf:

$$p(\tau_v) \propto \tau_v^{a_v-1} e^{-b_v\tau_v}. \quad (2.39)$$

Posterior Distribution

The joint posterior will be given by:

$$\begin{aligned} P(\beta, \Lambda_0, u, v \mid D) &\propto L(D \mid \beta, \Lambda_0, v) P(\beta) P(\Lambda_0) P(u \mid \tau_u) P(\tau_u) P(v \mid \tau_v) P(\tau_v) \\ &= \Pi_{i=1}^{n} \left[\Pi_{t \geq 0} [Y_i(t) \lambda_0(t) exp(\beta^T \mathbf{X}_i + u_j + v_j)]^{dN_i(t)} \right] \\ &\quad exp(- \left[Y_i(t) exp(\beta^T \mathbf{X}_i + u_j + v_j) d\Lambda_0(t) \right]) \\ &\times \exp \left[\frac{-1}{2} \left(\frac{\beta - \mu_\beta}{\sigma_\beta} \right)^2 \right] \\ &\times \Lambda_0^{(cd\Lambda_0^*(t) - 1)} e^{-c\Lambda_0(t)} \\ &\times \exp \left[\frac{-1}{2} u^2 \tau_u \right] \tau_u^{a_u - 1} e^{-b_u \tau_u} \\ &\times \exp \left[\frac{-\tau_v}{2} \sum_{i \sim j} (v_i - v_j)^2 \right] \tau_v^{a_v - 1} e^{-b_v \tau_v} \,. \end{aligned} \tag{2.40}$$

2.12 Accelerated Failure Time model (AFT) Model Specification

An accelerated failure time model is a model in which the effects of the covariates accelerate or decelerate the time to event of subjects. It can be used as a parametric alternative to the semi-parametric proportional hazards model and in instances where the proportional hazards assumption does not hold (65). The effect of the covariates is, therefore, multiplicative on the survival time and exponentiating the coefficients gives us the direct effect of the covariate on survival time (45). The AFT model is a log-linear regression model for time (T) to the event given by:

$$log(T) = \beta_0 + \beta \mathbf{X} + \sigma \epsilon \,, \tag{2.41}$$

where β_0 is the linear regression constant (the shape parameter), $\sigma\epsilon$ are the scaled residuals and $\beta\mathbf{X}$ is as defined above.

Bayesian AFT models have over the past been applied in many epidemiological studies some of which include cancer studies (44,45), nutrition studies (67), diabetes studies (65), to mention but a few and the implementation in OpenBugs is well demonstrated (46). We made use of the approach and methodological frameworks of these studies to fit our model. The AFT model is a log-linear model of the form:

$$log(T_{ij}) \quad = \quad \beta_0 + \boldsymbol{\beta}\mathbf{X}_i + u_j + v_j + \sigma\epsilon\,, \tag{2.42}$$

where T_{ij} is the survival time for individual i, $i = 1,...,n$ is the number of children, β are the regression coefficients, $\mathbf{X}_i$ are covariates, u_j are the unstructured spatial random effects and v_j are the structured spatial random effects for $j = 1,\ldots,m$ districts in Zimbabwe and ϵ are the random errors. The random erors can be used to find the probability density function(pdf) and survival function of the AFT model, if we make ϵ the subject of the formula,

$$\epsilon \quad = \quad \frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\,. \tag{2.43}$$

We can let $f_0()$, $S_0()$ and $h_0()$ be the respective baseline pdf, survival and hazards for our model and express our functions in a way similar to that of zhang et al (44).

Density Function

$$f(T_{ij} \mid u_j, v_j) \quad = \quad (\sigma T_{ij})^{-1} f_0(\epsilon) \tag{2.44}$$

$$\quad = \quad \frac{1}{\sigma T_{ij}} f_0 \left[\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right]. \tag{2.45}$$

Survival Function

$$S(T_{ij} \mid u_j, v_j) = S_0(\epsilon) \tag{2.46}$$

$$= S_0\left[\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right]. \tag{2.47}$$

Hazard Function

$$h(T_{ij} \mid u_j, v_j) = (\sigma T_{ij})^{-1} h_0(\epsilon) \tag{2.48}$$

$$= \frac{1}{\sigma T_{ij}} h_0\left[\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right]. \tag{2.49}$$

Likelihood Function

From equation 2.9, the likelihood can be expressed as a product of the density and survival functions, with censoring indicators.

$$L(T \mid \beta_0, \beta, \sigma, v) = \Pi_{i=1}^{n}\Pi_{j=1}^{m}\left[f(T_{ij})\right]^{\delta_i}\left[S(T_{ij})\right]^{1-\delta_i} \tag{2.50}$$

$$= \Pi_{i=1}^{n}\Pi_{j=1}^{m}\left[\frac{1}{\sigma T_{ij}} f_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]^{\delta_i} \left[S_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]^{1-\delta_i}. \tag{2.51}$$

Taking logs on both sides

$$log(L(T \mid \beta_0, \beta, \sigma, v)) = log\left\{\Pi_{i=1}^{n}\Pi_{j=1}^{m}\left[\frac{1}{\sigma T_{ij}} f_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]^{\delta_i} \left[S_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]^{1-\delta_i}\right\}. \tag{2.52}$$

Therefore :

$$\begin{aligned} l(T \mid \beta_0, \beta, \sigma, v) &= \sum_{i=1}^{n}\sum_{j=1}^{m}\left\{\delta_i log\left[\frac{1}{\sigma T_{ij}} f_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]\right. \\ &+ \left.(1-\delta_i) log\left[S_0\left(\frac{log(T_{ij}) - \beta_0 - \boldsymbol{\beta}\mathbf{X}_i - u_j - v_j}{\sigma}\right)\right]\right\}. \quad (2.53) \end{aligned}$$

The AFT model is a fully parametric model where the survival time is assigned a distribution and the three commonly used distributions are log-normal, log-logistic and Weibull distributions. To ensure that the log time follows these distributions, we may simply assign the normal, logistic and extreme value distributions to our baseline survival $S_0(.)$ and density $f_0(.)$ functions (68). The baseline pdf and survival functions for the selected model (log-normal) based on DICs (Appendix D) are $f_0 = \frac{1}{\sqrt{2\pi}} \exp(\epsilon_{ij}^2)$ and $S_0 = 1 - \phi(\epsilon_{ij})$. We can substitute f_0 and S_0 into our likelihood $l(T \mid \beta_0, \beta, \sigma, v)$.

Prior Specifications

The prior distributions for our parameters are expressed in the following way.

Regression coefficients (β) and shape parameter (β_0) follow a gaussian distribution such that; $\beta \sim N(\mu, \sigma^2)$ and $\beta_0 \sim N(\mu, \sigma^2)$ and they can be expressed together at the same time as: $\boldsymbol{\beta} \sim N(\mu_\beta, \sigma_\beta^2)$

$$\Rightarrow p(\beta) = \frac{1}{\sqrt{2\pi\sigma_\beta^2}} \exp\left[\frac{-1}{2}\left(\frac{\beta - \mu_\beta}{\sigma_\beta}\right)^2\right] \quad (2.54)$$

$$\propto \exp\left[\frac{-1}{2}\left(\frac{\beta - \mu_\beta}{\sigma_\beta}\right)^2\right]. \quad (2.55)$$

The Scale parameter (σ) was assigned a Gamma prior; $\sigma \sim \Gamma(a,b)$ with pdf given by:

$$\begin{aligned} p(\sigma) &= \frac{b^a}{\Gamma(a)}\sigma^{a-1}e^{-b\sigma} && (2.56)\\ &\propto \sigma^{a-1}e^{-b\sigma}\,. && (2.57)\end{aligned}$$

The prior specification of the spatial random effects in the AFT models were specified in the same way as those in the Cox-PH models. The unstructured random effects (u) were assumed to follow a Gaussian distribution with mean = 0 and precision τ_u; u $N(0.0,\tau_u)$ (Equation 2.25) with the precision parameter assigned a Gamma hyper-prior (Equation 2.38). The structured spatial effects v were assumed to follow an ICAR prior (Equation 2.35) with the corresponding hyper-prior assumed to follow a Gamma distribution (Equation 2.39).

Posterior Distribution

The joint posterior will be given by:

$$\begin{aligned} P(\beta_0,\beta,\sigma,v \mid T) &\propto l(T\mid,\beta,\sigma,u,v)P(\beta)P(\sigma)P(u\mid\tau_u)P(\tau_u)P(v\mid\tau_v)P(\tau_v)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}\left\{\delta_i log\left[\frac{1}{\sigma T_{ij}}f_0\left(\frac{log(T_{ij})-\beta_0-\boldsymbol{\beta}\mathbf{X}_i-u_j-v_j}{\sigma}\right)\right]\right.\\ &+ \left.(1-\delta_i)log\left[S_0\left(\frac{log(T_{ij})-\beta_0-\boldsymbol{\beta}\mathbf{X}_i-u_j-v_j}{\sigma}\right)\right]\right\}\\ &\times \exp\left[\frac{-1}{2}\left(\frac{\beta-\mu_\beta}{\sigma_\beta}\right)^2\right]\times\sigma^{a-1}e^{-b\sigma}\\ &\times \exp\left[\frac{-1}{2}u^2\tau_u\right]\tau_u^{a_u-1}e^{-b_u\tau_u}\\ &\times \exp\left[\frac{-\tau_v}{2}\sum_{i\sim j}(v_i-v_j)^2\right]\tau_v^{a_v-1}e^{-b_v\tau_v}\,. \qquad (2.58)\end{aligned}$$

2.13 Model Fitting in R and OpenBugs

OpenBugs, a free Bayesian statistical analysis software package and *spBayesSurv package* in R statistical software were used to run Cox PH spatial models and the AFT spatial models. OpenBugs uses the Gibbs sampling technique for its MCMC sampler while BE in R uses the Metropolis-Hastings algorithm for sampling (47). A total of 21,000 iterations was set with a burn-in period of 1000 and thinning after every 5 iterations. We ran two chains for the models while utilizing a less informative hyperprior ($\tau \sim \Gamma(0.001, 0.001)$) and a more informative hyperprior ($\tau \sim \Gamma(0.1, 0.1)$) for the precision of the spatial prior. The model codes used in the analysis are provide in Appendix E.

2.14 Model Selection

To compare non-spatial models, we made use of the Akaike Information Criterion (AIC) which has a penalty for more complex models and is calculated from the maximum likelihood estimate (MLE) of the models. The model with the lowest AIC value is selected as the better fitting model (69). It is given by given by:

$$AIC \quad = \quad -2\log(\text{likelihood}) + 2k\,, \tag{2.59}$$

where k is the number of variables in the model.

For comparison of Bayesian models, Deviance Information Criterion (DIC) which is a Bayesian version of AIC was used and the model with the lowest DIC value was regarded as the better model. In calculating the DIC, there is a penalty imposed for complex models which are done to improve goodness of fit and the more complex

models have a higher penalty. The formula for finding the DIC may be written as:

$$DIC = \bar{D} + pD\,, \tag{2.60}$$

where $\bar{D} = -2$log-likelihood and pD is the effective number of parameters in the model (69).

2.15 Model Assessment and Goodness of Fit

Model goodness of fit is how well the model used fits our data and how well the outcome is described by the model (70). The goodness of fit was assessed by observing the mixing and convergence of time series plots, the tailing off and convergence of autocorrelation plots and how well the kernel density plots produced the graphs of the probability distribution for the variables. The model diagnostics plots for the fitted R and OpenBugs BE models are shared in Appendix F. To assess the PH assumption, we based our tests on the Schoenfeld residuals and the global test (p = 0.038) gave evidence of a violation of the PH assumption for the 2011 model while for 2015 Cox-PH regression models (p =0.35), there was no evidence of a violation of the PH assumption. Fortunately, the AFT models can be used as an alternative in situations where the PH assumption does hold and the use of both Cox-PH and AFT models can give more accurate results.

2.16 Ethical Considerations

The primary ZDHS surveys were cleared by the relevant ethics committees and proper protocols were followed for the studies not to pose any harm to human participants. Permission to use the ZDHS data was sought and granted from the Measure DHS and these data are publicly available. The University of Witwatersrand Human Research

Ethics Committee (medical) granted ethical clearance to proceed with this research project, reference number W-CBP-200904-01 (Appendix G).

Chapter 3

RESULTS

3.1 Introduction

In this chapter, we present the results of U5M, defined as the death of a child before its fifth birthday. A survival analysis approach was taken, and several Cox-PH and AFT survival models were fitted using R and OpenBugs software. We first report results on descriptive statistics stratified by mortality, after which descriptive spatial distribution maps which show the U5M rates across the country are displayed. The subsequent section is on results from univariate survival analysis presented in the form of a rate per 1000 person-years accompanied by log-rank test and Kaplan-Meier plots to examine the significance of variables in the univariate analysis. This is followed by the section on results from adjusted non-spatial Cox PH and AFT survival models which are extended by incorporating spatial random effects to get the corresponding spatial models. Lastly, we show the spatial distribution of under-five mortality in Zimbabwe for the selected Bayesian mixed-effects models.

3.2 Descriptive Statistics

The descriptive characteristics of the participants that influence U5M are summarised in Table 3.1 below. The results were stratified by the mortality outcome for the two years considered in this study. Bivariate analysis was determined for each variable, for each year to determine any significant differences.The results in Table 3.1 accounted for sampling weights of the ZDHS surveys.

Results from our analysis suggest that in the 2011 survey, 6.93% of the children died before turning five years of age. There were more single births (97.30%, n=5,445) than multiple births (2.70%, n=151). There was a fair balance in male (50.35%, n=2817) and female (49.65%, n=2,779) sub-populations in 2011. Majority of the babies had an average size at birth (47.82%, n=2,678). The average age of the mothers was 27.5 (SD=6.33) years. Majority of the mothers had attained secondary/higher level of education (65.89%, n=3,687), were married or in unions (86.21%, n=4,827) and were using contraceptives (62.39%, n=3,491). Most mothers fell in the poor wealth index (43.88%, n=2,455). Majority of the participants were from rural areas (70.23%, n=3,930) and were of the Apostolic religion (45.71%, n=2,558). There was a significant association between U5M and type of birth (p-value <0.001), sex of child (p-value=0.005), marital status (p-value <0.001), contraceptive use (p-value <0.001) and religion (p-value=0.06).

Table 3.1: Baseline Characteristics of study participants, for 2011 and 2015 data, stratified by mortality.

Variable	Category	2011				2015			
		Overall n (%) 5596(100)	**Dead n (%) 388(6.93)**	**Alive n (%) 5208(93.07)**	**p-value**	**Overall n (%) 6418 (100)**	**Dead n (%) 363 (5.65)**	**Alive n (%) 6055(94.35)**	**p-value**
Type of birth	Single	5445(97.30)	346 (6.35)	5099(93.75)	<0.001*	6191(96.47)	324(5.24)	5867(94.76)	<0.001*
	Multiple	151(2.70)	42 (27.83)	109(72.17)		227(3.53)	38(16.82)	188(83.18)	
Sex of child	Male	281(50.35)	225 (7.99)	2592(92.01)	0.005*	3148(49.05)	198(6.30)	2950(93.70)	0.075
	Female	2779(49.65)	163 (5.86)	2616(94.14)		3270(50.95)	165(5.03)	3105(94.97)	
Size of child at birth	Small	709(12.67)	86 (12.14)	623(87.86)	<0.001*	954(14.86)	71(7.45)	883(92.55)	0.0056*
	Average	2678(47.86)	162(6.04)	2516(93.96)		3129(48.75)	139(4.44)	2990(95.56)	
	Large	2209(39.47)	140(6.34)	2069(93.66)		2335(36.39)	153(6.54)	2182(93.46)	
Mother's age at childbirth in years	Mean±SD	27.51±6.33	28.10±6.18	27.47±6.34	0.11	28.67±6.53	28.35±6.52	28.69±6.53	0.49
Mother's education level	No Education	95(1.69)	6(6.13)	89(93.87)	0.30	76(1.18)	6(8.16)	70(91.84)	<0.001*
	Primary	1814(32.42)	141(7.79)	1673(92.21)		2038(31.76)	155(7.60)	1883(92.40)	
	Secondary/Higher	3687(65.89)	241(6.53)	3446(93.47)		4304(67.06)	202(4.69)	4102(95.31)	
Marital status	Single	200(3.57)	5(2.26)	195(97.74)	0.001*	251(3.91)	7(2.61)	244(97.39)	0.021*
	Married	4824(86.21)	321(6.66)	4503(93.34)		5531(86.18)	307(5.56)	5224(94.44)	
	Widowed/Divorced	572(10.22)	62(10.84)	510(89.16)		636(9.91)	49(7.68)	587(92.32)	
Contraceptive use	No	2104(37.61)	211(10.02)	1894(89.98)	<0.001*	1852(28.86)	163(8.81)	1689(91.19)	<0.001*
	Yes	3491(62.39)	177(5.07)	3314(94.93)		4366(71.14)	200(4.37)	4366(95.63)	
Wealth Index	Poor	2455(43.88)	179(7.31)	2276(92.69)	0.67	2729(42.52)	169(6.21)	2559(93.79)	0.0038*
	Middle	1070(19.12)	73(6.81)	997(93.19)		1098(17.10)	82(7.46)	1016(92.54)	
	Rich	2071(37.00)	136(6.55)	1935(93.45)		2592(40.38)	112(4.30)	2480(95.70)	
Residence	Urban	1666(29.77)	118(7.08)	1548(92.92)	0.83	2027(31.58)	90(4.42)	1937(95.58)	0.017*
	Rural	3930(70.23)	270(6.87)	3660(93.13)		4391(68.42)	273(6.22)	4118(93.78)	
Religion	Apostolic	2558(45.71)	207(8.09)	2351(91.91)		3228(50.29)	229(7.11)	2998(92.89)	<0.001*
	Christian	1507(26.93)	82(5.47)	1425(94.53)	0.06*	1376(21.43)	70(5.08)	1306(94.92)	
	Pentecostal	1029(18.38)	60(5.84)	969(94.16)		1356(21.13)	46(3.36)	1311(96.64)	
	Other	502(8.97)	39(7.65)	463(92.35)		458(7.14)	18(3.90)	440(96.10)	

* $p<0.05$, which is significant at a 5% significance level.
Estimates accounted for ZDHS sampling weights

The results from the 2015 survey, suggested that 5.65% of the children died before five years. Like in 2011, there were more single births (96.47%, n=6,191) than multiple births (3.53%, n=227). There were more boys (50.95%, n=3,148) than girls. The majority of the babies were averagely sized (48.74%, n=3,129). Most mothers had secondary or higher education (67.06%, n=4,304), were married (86.18%, n=5,531) and were using some form of a contraceptive method (71.14%, n=4,366). There were more participants from the rural areas (68.42%, n=4,391) and most were of the Apostolic religion (50.29%, n=3,228). There was a significant association between U5M and type of birth (p-value <0.001), size of the baby (p-value=0.0056), level of education of the mother (p-value <0.001), contraceptive use (p-value <0.001), marital status (p-value=0.021), wealth index (p-value=0.0038), place of residence (p-value=0.017) and religion (p-value <0.001). Comparing the U5M in 2011 and 2015 using the Z proportion test between various baseline characteristics, some sub-groups showed an increase in mortality like Apostolic religion (53.38% in 2011 vs 63.25% in 2015) and this was statistically significant (p-value =0.041). In contrast some variables showed a decrease in U5M like contraceptive use (54.35% in 2011 vs 44.98% in 2015, p-value= 0.05) and other religions (9.90% in 2011 vs 4.93% in 2015, p-value=0.078).

3.3 Mortality Rate: Standardised Mortality Ratio and KM Curves

3.3.1 Standardised Mortality Ratio (SMR)

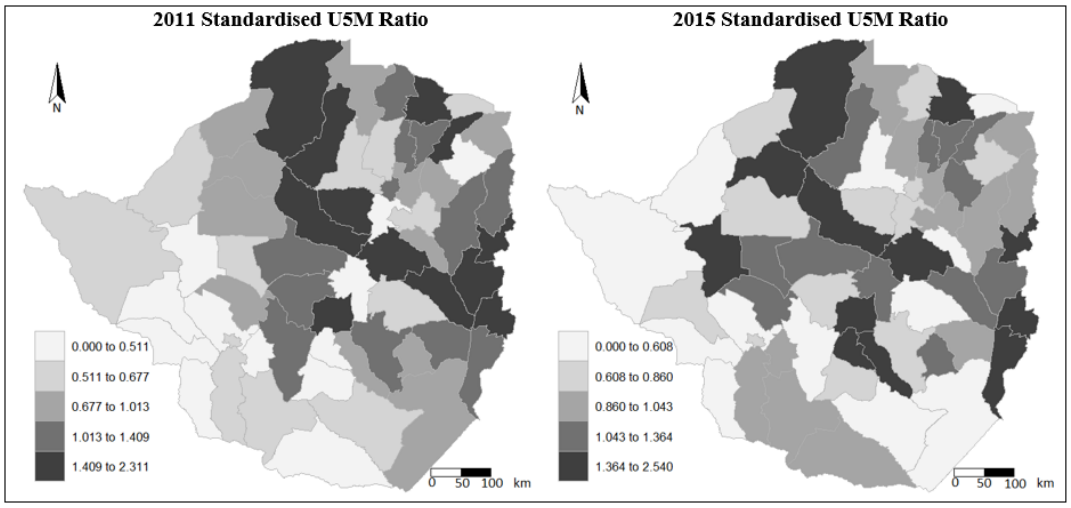


Figure 3.1: Maps showing the standardised U5M ratio of the districts in Zimbabwe in 2011 (left) and 2015(right) ZDHS.

The standardised mortality ratios for U5M across the country's districts are shown in Figure 3.1. The light colours denote lower mortality risks while the dark colours denote higher U5M risks. The U5M risks in 2011 were higher than the average U5M risk, in the northern parts of the country, stretching down through the middle parts to the Eastern parts of the country and lower in the Western and Southern regions. The 2015 map shows a similar trend to the 2011 map, with higher risks of mortality in the Northern, Central and Eastern districts of the country.

3.3.2 Mortality Rate and KM Curves

The overall mortality rate was 32.53 per 1000 person-years in 2011 compared to 23.66 per 1000 person-years in 2015 and this difference was statistically significant, p-value=0.002. Figure 3.2 is the KM curve for the overall U5M rates for 2011 and

2015 survey data.

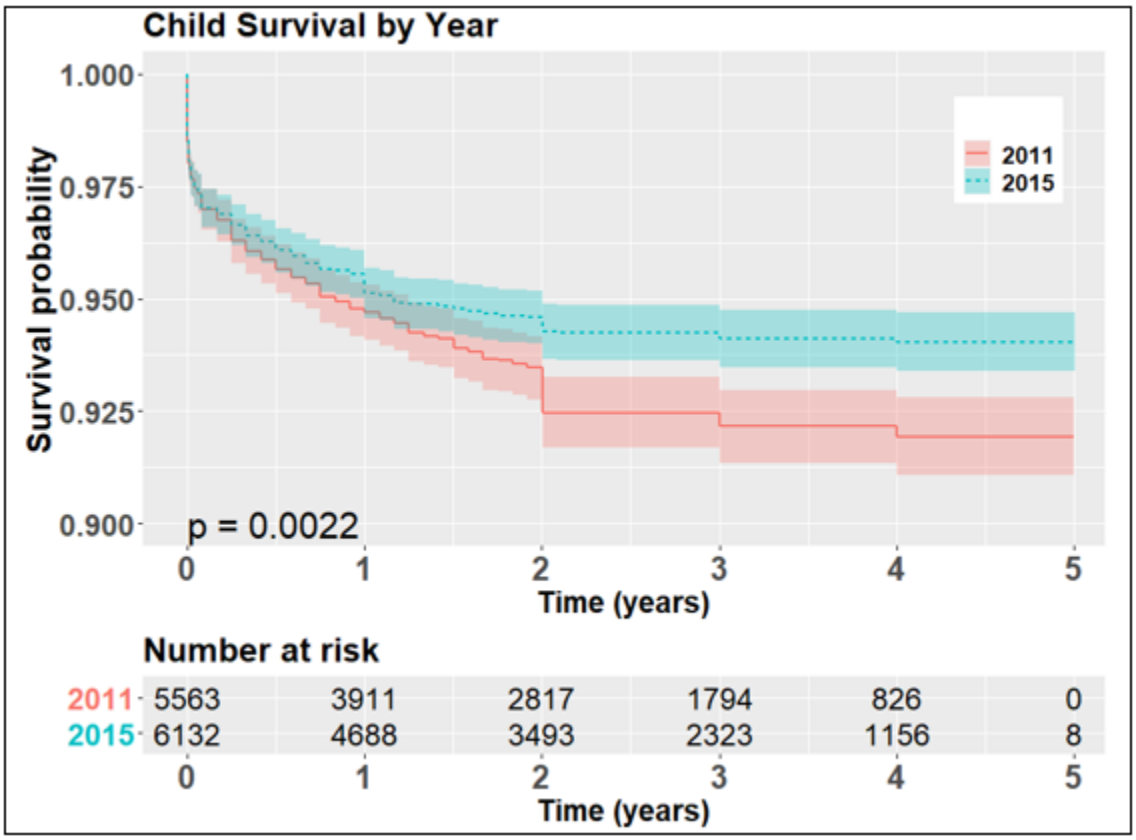


Figure 3.2: The overall Kaplan-Meier curve for the U5M rates for 2011 and 2015.

Table 3.2: Mortality rate per 1000 person-years and log-rank test of under-five children in Zimbabwe for 2011 and 2015 ZDHS data.

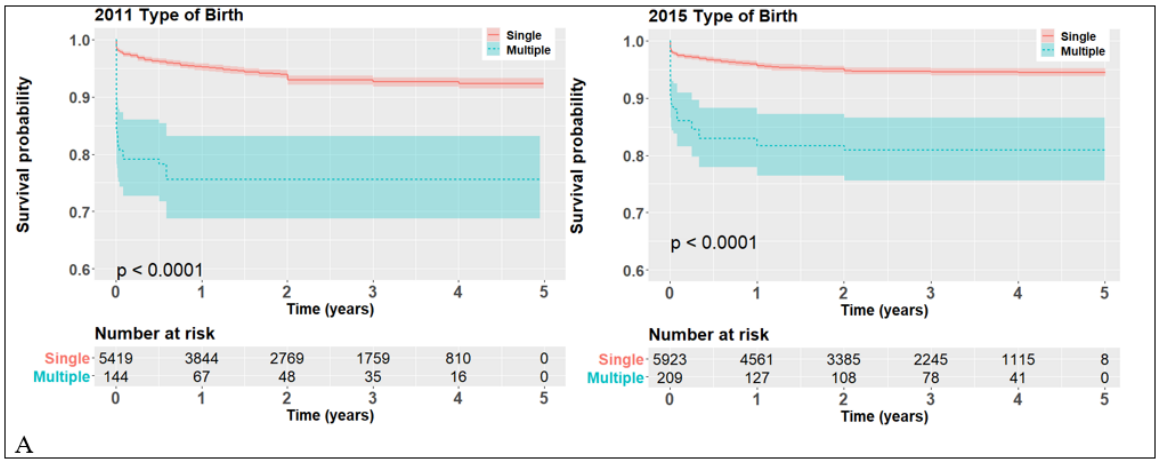
Variable	Category	2011			2015		
		Rate per 1000 person-years	**95% CI**	**Log-rank p-value**	**Rate per 1000 person-years**	**95% CI**	**Log-rank p-value**
Type of birth	Single	29.58	26.24-33.45	<0.001*	21.87	19.20-25.02	<0.001*
	Multiple	180.32	117.24-282.14		78.10	51.74-120.58	
Sex of child	Male	38.23	32.82-44.75	0.015*	26.99	22.82-32.13	0.086
	Female	26.97	22.73-32.23		20.59	17.19-24.86	
Size of child at birth	Small	61.42	47.38-80.69	<0.001*	33.16	25.40-43.97	<0.001*
	Average	28.02	23.64-33.44		18.39	15.11-22.60	
	Large	29.49	24.38-35.95		27.11	22.33-33.22	
Mother's education level	No Education	23.31	9.70-67.99	0.31	36.80	16.10-97.54	<0.001*
	Primary	36.12	30.04-43.76		32.28	26.72-39.32	
	Secondary/Higher	31.01	26.78-36.09		19.46	16.50-23.10	
Marital status	Single	11.96	4.35-44.95	<0.001*	13.53	6.31-33.99	0.005*
	Married	31.46	27.76-35.79		23.26	20.33-26.74	
	Widowed/Divorced	46.55	34.86-63.28		29.86	21.61-42.27	
Contraceptive use	No	45.24	38.64-53.26	<0.001*	37.57	31.13-45.70	<0.001*
	Yes	24.38	20.66-28.95		18.16	15.43-21.53	
Wealth Index	Poor	34.27	29.07-40.64	0.51	27.04	22.62-32.55	0.0034*
	Middle	32.16	24.66 - 42.62		32.26	24.86-42.53	
	Rich	30.66	25.24-37.59		17.08	13.69-21.58	
Residence	Urban	32.99	26.56-41.42	0.70	17.51	13.64-22.82	0.0054*
	Rural	32.33	28.29-37.10		26.74	23.22-30.94	
Religion	Apostolic	38.37	32.91-44.96		30.15	25.80-35.42	
	Christian	25.61	20.13-33.05	0.032*	21.23	16.06-28.59	<0.001*
	Pentecostal	27.43	20.42-37.62		13.82	9.99-19.63	
	Other	34.22	23.14-52.55		15.78	8.74-31.30	

* $p<0.05$, which is significant at a 5% significance level.

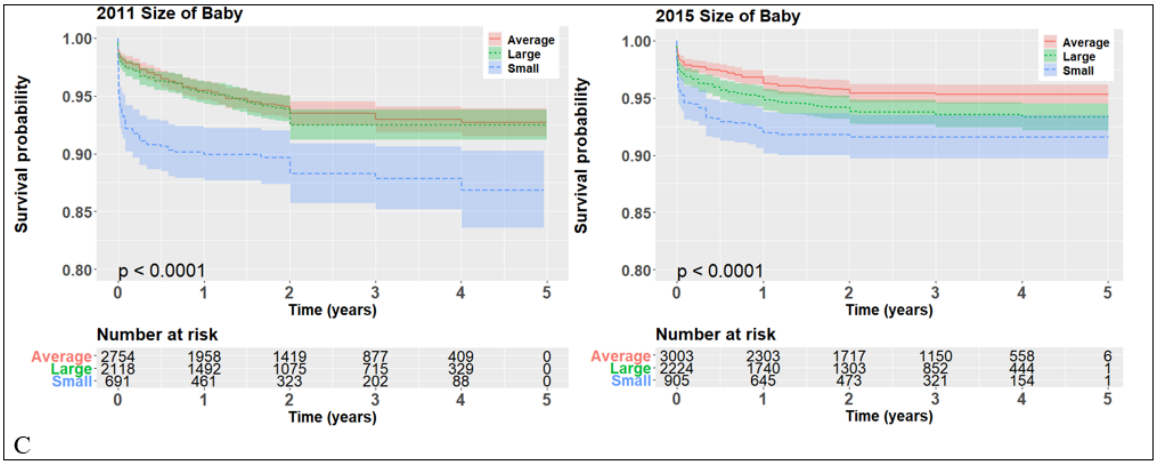
The mortality rates in 2011 were significantly higher than those of 2015 (p-value=0.0022). In 2011, the mortality rates were significantly higher (p-value <0.001) in twins (180.32 per 1000 person-years) than in single births (29.58 per 1000 person-years) (Figure 3.3 panel A); in males (38.23 per 1000 person-years) than in females (26.96 per 1000 person-years) and this was statistically significant (p-value=0.015) (Figure 3.3 panel B); in small size children at birth (61.42 per 1000 person-years) and large children at birth (29.49 per 1000 person-years) compared to averaged sized children at birth (28.02 per 1000 person-years) and this was statistically significant (p-value <0.001) (Figure 3.3 panel C). Child mortality rates were substantially high among widowed and divorced mothers (46.55 per 1000 person-years) (Figure 3.3 panel D), among mothers who had attained primary level of education (36.12 per 1000 person-years), among mothers who did not follow any contraceptive method (45.24 per 1000 person-years) (Figure 3.3 panel E) and among those who were in Apostolic religion (38.37 per 1000 person-years) and these differences were all statistically significant between groups (p-value<0.05).

For the 2015 survey, the mortality rates were higher in twins/multiple births (78.1 per 1000 person-years) than in single births (21.87) and this was statistically significant (p-value<0.001) (Figure 3.3 panel A); in males (26.99 per 1000 person-years) than in females (20.59 per 1000 person-years) and this was marginally significant (p-value=0.086) (Figure 3.3 panel B); in small size children at birth (33.16 per 1000 person-years) and large children at birth (27.11 per 1000 person-years) compared to averaged sized children at birth (18.39 per 1000 person-years) and this was statistically significant (p-value<0.001) (Figure 3.3 panel C). Besides, the mortality rates were significantly high among widowed and divorced mothers (29.86 per 1000 person-years) (Figure 3.3 panel D), among mothers who had no education (36.8 per 1000 person-years), among mothers who did not use contraceptives (37.57 per 1000 person-years)

(Figure 3.3 panel E) and among those who were in Apostolic religion (30.15 per 1000 person-years) and these differences were all statistically significant between groups (p-value<0.05).







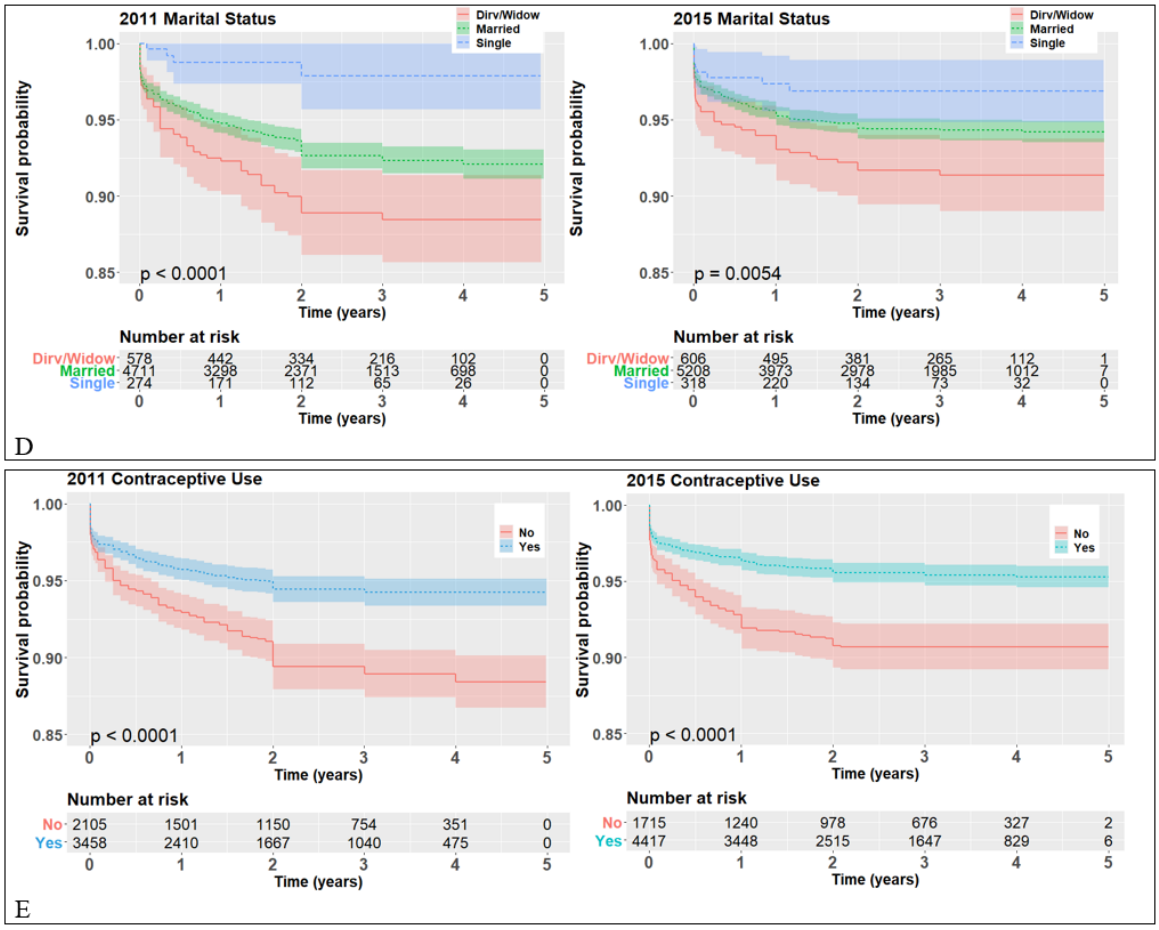


Figure 3.3: Kaplan-Meier curves for child survival in Zimbabwe by different characteristics.

3.4 2011 Regression Models

We fitted non-spatial MLE survival models and compared them to corresponding Bayesian spatial models (Cox-PH and AFT) to establish the benefit of including the spatial random effects in the models. Results for the spatial and non-spatial regression models for 2011 survey data are shown in Table 3.3. The selected models were model 3 from the Cox-PH models and model 6 from the AFT models.

Table 3.3: Adjusted non-spatial and spatial Cox-PH and AFT models for 2011 ZDHS data.

Variable	**Category**	**2011 Adjusted Cox-PH Models**			**2011 Adjusted AFT Models**		
		MLE Non Spatial Model	**Bayesian spatial Models**		**MLE Non Spatial Model**	**Bayesian spatial Models**	
		Cox-PH	**Cox-PH in R**	**Cox-PH in OpenBugs**	**AFT**	**AFT in R**	**AFT in OpenBugs**
		HR[95% CI] model 1	**PHR[95% CI] model 2**	**PHR[95% CI] model 3**	**HR[95% CI] model 4**	**PTR[95% CI] model 5**	**PTR[95% CI] model 6**
Type of birth	Single	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Multiple	**4.03[2.76-5.88]**	**3.90[2.78-5.47]**	**3.84[2.56-5.50]**	**0.0048[0.0011-0.022]**	**0.0065[0.0011-0.036]**	**0.0057[0.0012-0.027]**
Sex of child	Male	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Female	**0.75[0.60-0.92]**	**0.76[0.68-0.84]**	**0.74[0.60-0.90]**	**2.42[1.24-4.73]**	**2.59[1.25-5.31]**	**2.59[1.32-5.25]**
Size of child at birth	Small	**1.74[1.31-2.31]**	**1.74[1.42-2.08]**	**1.73[1.26-2.25]**	**0.12[0.046-0.32]**	**0.12[0.048-0.35]**	**0.12[0.042-0.31]**
	Average	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Large	1.06[0.84-1.34]	0.98[0.83-1.16]	**0.99[0.77-1.25]**	0.76[0.37-1.59]	0.98[0.44-2.09]	0.95[0.44-2.07]
Mother's age at childbirth in years		0.99[0.91-1.08]	1.00[0.94-1.06]	1.00[0.92-1.08]	0.99[0.76-1.28]	0.97[0.75-1.26]	0.97[0.74-1.28]
Mother's education level	No Education	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Primary	1.27[0.55-2.90]	**1.49[1.02-2.20]**	1.66[0.70-3.43]	0.48[0.036-6.25]	0.33[0.026-4.10]	0.25[0.017-3.14]
	Secondary/Higher	1.15[0.50-2.64]	1.34[0.88-2.09]	1.48[0.62-3.08]	0.60[0.046-7.97]	0.48[0.038-5.67]	0.35[0.024-4.14]
Marital status	Single	**0.23[0.084-0.61]**	**0.31[0.13-0.64]**	**0.32[0.083-0.74]**	**93.55[6.47-135.35]**	**55.41[3.16-1262.43]**	**54.27[3.65-1227.83]**
	Married	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Widowed/Divorced	1.23[0.92-1.65]	1.24[0.94-1.60]	1.28[0.96-1.68]	0.54[0.20-1.47]	0.50[0.17-1.47]	0.47[0.16-1.35]
Contraceptive use	No	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Yes	**0.55[0.44-0.68]**	**0.52[0.43-0.63]**	**0.53[0.42-0.66]**	**5.79[2.84-11.80]**	**6.81[2.98-15.34]**	**6.63[3.14-14.28]**

Variable	Category	2011 Adjusted Cox-PH Models			2011 Adjusted AFT Models		
		MLE Non Spatial Model	**Bayesian spatial Models**		**MLE Non Spatial Model**	**Bayesian spatial Models**	
		Cox-PH	**Cox-PH in R**	**Cox-PH in OpenBugs**	**AFT**	**AFT in R**	**AFT in OpenBugs**
		HR[95% CI] model 1	**PHR[95% CI] model 2**	**PHR[95% CI] model 3**	**HR[95% CI] model 4**	**PTR[95% CI] model 5**	**PTR[95% CI] model 6**
Wealth Index	Poor	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Middle	1.14[0.85-1.52]	1.10[0.88-1.38]	1.10[0.81-1.45]	0.70[0.28-1.76]	0.76[0.33-1.80]	0.78[0.29-2.11]
	Rich	0.86[0.60-1.23]	0.86[0.67-1.13]	0.87[0.60 -1.22]	1.73[0.58-5.21]	1.84[0.62-6.03]	1.95[0.65-6.18]
Residence	Urban	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Rural	0.72[0.51-1.02]	**0.68[0.57-0.80]**	**0.70[0.48-1.00]**	2.65[0.91-7.75]	3.84[1.01-12.28]	**3.59[1.09-12.60]**
Religion	Apostolic	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Christian	**0.70[0.54-0.91]**	**0.79[0.62-0.99]**	0.79[0.59-1.03]	**2.69[1.17-6.17]**	1.98[0.89-4.74]	1.90[0.77-4.76]
	Pentecostal	**0.69[0.50-0.94]**	**0.69[0.54-0.90]**	**0.72[0.51-0.94]**	**3.01[1.11-8.17]**	**3.05[1.19-7.87]**	**2.88[1.06-7.98]**
	Other	0.74[0.50-1.08]	**0.74[0.58-0.94]**	0.76[0.50-1.08]	2.44[0.73-8.23]	2.42[0.65-8.57]	2.32[0.70-8.38]
	AIC	**5945**			**6602.9**		
	DIC (pD)		**6577 (35.04)**	**4809 (75.34)**		**6570 (35.93)**	**6562 (17.25)**

Bold estimates: significant at a 5% significance level.
PHR - Posterior Hazards Ratio; PTR - Posterior Time Ratio.
CI for MLE - 95%confidence interval; CI for Bayesian model - 95% credible interval.

With reference to the spatial adjusted Bayesian Cox-PH model for the 2011 survey and adjusting for other factors, children born through multiple births had an increased risk of mortality compared to those born through singleton births (posterior HR (PHR) = 3.84; 95% CI: 2.56 - 5.50). Being a girl was protective from death as girls were 26% less likely to die than boys (PHR = 0.74; 95% CI: 0.60 - 0.90). With reference to normal-sized babies at birth, children who were born underweight had a 73% risk of dying before five years (PHR = 1.73; 95% CI: 1.26 - 2.25). Being born to a single mother was protective since a child born by a single mother was 68% less likely to die compared to a child from a married mother (PHR = 0.32; 95% CI: 0.083 - 0.74). It is however important to note that there were a few single mothers who participated in the survey. Contraceptive use proved to be a critical predictor of child mortality. Children born to mothers who were using contraceptives were 47% less likely to die before the age of five compared to those whose mothers were not using contraceptives (PHR = 0.53; 95% CI: 0.42 - 0.66). Place of residence was marginally a significant factor of U5M. Rural residence children were 30% less likely to die compared to urban counterparts (PHR = 0.70; 95% CI: 0.48 - 1.00). Religious beliefs and practices appeared to influence a child's survival and being of the Pentecostal faith was protective of death compared to Apostolic faith (PHR = 0.72; 95% CI: 0.51 - 0.94).

With reference to the spatial adjusted Bayesian AFT model for the 2011 survey and adjusting for other factors, children born through multiple births lived for a significantly shorter time when compared to children from singular births. On average, children born through multiple births survived for 0.57% of the time lived by children from singular births as most of them died during the early days of life (PTR = 0.0057; 95% CI: 0.0012- 0.027). Being a boy accelerated the time to death as girls survived for more than double the time survived by boys (PTR = 2.59; 95% CI: 1.32- 5.25). Child birthweight was an effective predictor of survival and in 2011, underweight

children survived for just 12% of the average time for children with normal weight (PTR = 0.12; 95% CI: 0.042- 0.31). Single mothers (PTR = 54.27; 95% CI: 3.65, 1227.83) and the children whose mothers practised family planning (PTR = 6.63; 95% CI: 3.14, 14.28) survived longer; however, the credible interval was less precise for single mothers. Being from rural areas (TR = 3.59; 95% CI: 1.09, 12.60) and being in Pentecostalism (PTR = 2.88; 95% CI: 1.06, 7.98) significantly decelerated the time to death of a child.

3.5 2015 Regression Models

Results for the spatial and non-spatial regression models for 2015 survey data are shown in Table 3.4. The selected models were model 9 from the Cox-PH models and model 11 from the AFT models.

Table 3.4: Adjusted non-spatial and spatial Cox-PH and AFT models for 2015 ZDHS data.

Variable	Category	2015 Adjusted Cox-PH Models			2015 Adjusted AFT Models		
		MLE Non Spatial Model	Bayesian spatial Models		MLE Non Spatial Model	Bayesian spatial Models	
		Cox-PH HR[95% CI] model 7	Cox-PH in R PHR[95% CI] model 8	Cox-PH in OpenBugs PHR[95% CI] model 9	AFT HR[95% CI] model 10	AFT in R PTR[95% CI] model 11	AFT in OpenBugs PTR[95% CI] model 12
Type of birth	Single Multiple	1(reference) **4.49[3.15-6.41]**	1(reference) **4.45[3.25-6.10]**	1(reference) **4.52[3.11-6.28]**	1(reference) **0.0024[0.00042-0.012]**	1(reference) **0.0019[0.00036-0.011]**	1(reference) **0.0019[0.00033-0.011]**
Sex of child	Male Female	1(reference) **0.79[0.63-0.98]**	1(reference) **0.77[0.65-0.91]**	1(reference) **0.79[0.63-0.96]**	1(reference) **2.55[1.14-5.70]**	1(reference) **2.76[1.07-6.25]**	1(reference) **2.65[1.15-6.32]**
Size of child at birth	Small Average Large	**1.81[1.35-2.44]** **1(reference)** 1.46[1.14-1.88]	**1.83[1.57-2.15]** 1(reference) **1.48[1.28-1.68]**	**1.81[1.32-2.37]** 1(reference) **1.47[1.15-1.85]**	**0.099[0.032-0.30]** 1(reference) **0.24[0.10-0.60]**	**0.080[0.024-0.24]** 1(reference) **0.21[0.077-0.54]**	**0.091[0.029-0.30]** 1(reference) **0.23[0.094-0.57]**
Mother's age at childbirth in years		0.93[0.85-1.01]	**0.92[0.87-0.97]**	**0.93[0.85-1.00]**	1.33[0.98-1.80]	**1.35[1.01-1.78]**	**1.36[1.02-1.83]**
Mother's education level	No Education Primary Secondary/Higher	1(reference) 0.62[0.29-1.34] **0.45[0.21-0.97]**	1(reference) **0.64[0.40-0.99]** **0.46[0.29-0.71]**	1(reference) 0.69[0.32-1.42] 0.49[0.23-1.02]	1(reference) 3.50[0.14-89.70] 10.96[0.43-282.10]	1(reference) 3.67[0.09-93.99] 11.92[0.29-309.32]	1(reference) 4.19[0.13-9.44] 13.39[0.44-317.67]
Marital status	Single Married Widowed/Divorced	0.52[0.26-1.02] 1(reference) 1.30[0.95-1.78]	**0.54[0.33-0.85]** 1(reference) **1.30[1.16-1.45]**	**0.55[0.24-0.99]** 1(reference) 1.31[0.93-1.76]	8.13[0.89-74.37] 1(reference) 0.39[0.10-1.13]	**10.23[1.10-91.96]** 1(reference) 0.31[0.088 - 1.20]	9.13[0.94-103.44] 1(reference) 0.32[0.094-1.14]
Contraceptive use	No Yes	1(reference) 0.51[0.40-0.63]	1(reference) **0.50[0.42-0.60]**	1(reference) **0.51[0.41-0.64]**	1(reference) **10.93[4.59-26.00]**	1(reference) **11.84[4.41-29.20]**	1(reference) **12.06[4.88-29.61]**

Variable	Category	2015 Adjusted Cox-PH Models			2015 Adjusted AFT Models		
		MLE Non Spatial Model	Bayesian spatial Models		MLE Non Spatial Model	Bayesian spatial Models	
		Cox-PH HR[95% CI] model 7	Cox-PH in R PHR[95% CI] model 8	Cox-PH in OpenBugs PHR[95% CI] model 9	AFT HR[95% CI] model 10	AFT in R PTR[95% CI] model 11	AFT in OpenBugs PTR[95% CI] model 12
Wealth Index	Poor	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Middle	1.27[0.94 -1.72]	1.25[1.06-1.49]	1.28[0.96-1.72]	0.50[0.16-1.57]	0.46[0.12-1.58]	0.49[0.14-1.70]
	Rich	0.95[0.60-1.48]	0.93[0.65-1.36]	0.94[0.57-1.41]	1.16[0.24-5.69]	1.14[0.21-6.56]	1.30[0.22-7.89]
Residence	Urban	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Rural	0.98[0.63-1.54]	0.95[0.63-1.38]	0.96[0.57-1.46]	0.96[0.20-4.62]	0.95[0.17-6.01]	1.10[0.20-6.01]
Religion	Apostolic	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)	1(reference)
	Christian	0.83[0.62-1.10]	**0.83[0.71-0.99]**	0.84[0.63-1.12]	1.67[0.59-4.71]	1.83[0.62-5.19]	1.69[0.56-5.15]
	Pentecostal	**0.60[0.43-0.83]**	**0.58[0.49-0.68]**	**0.61[0.42-0.84]**	**4.79[1.53-14.97]**	**5.41[1.48-18.86]**	**5.24[1.68-17.46]**
	Other	**0.46[0.26-0.81]**	**0.44[0.30-0.64]**	**0.46[0.24-0.78]**	**12.46[1.89-82.31]**	**16.27[2.59-105.84]**	**16.07[2.23-138.38]**
	AIC	**5459**			**5897**		
	DIC (pD)		**5878 (28.25)**	**4552 (91.96)**		**5892 (19.40)**	**5899 (18.09)**

Bold estimates: significant at a 5% significance level.
PHR - Posterior Hazards Ratio; PTR - Posterior Time Ratio.
CI for MLE - 95%confidence interval; CI for Bayesian model - 95% credible interval.

With reference to the spatial adjusted Bayesian Cox-PH model for the 2015 survey and adjusting for other factors, multiple births children were 4.52 times more likely to die compared to singular births (PHR = 4.52; 95% CI: 3.11 - 6.28). Survival amongst girls was significantly higher than in boys as girls were 21% less likely to die (PHR = 0.79; 95% CI: 0.63 - 0.96). The size of the baby at birth was an important predictor of death. Children born underweight were 1.81 times (PHR = 1.81; 95% CI: 1.32 - 2.37) and those children born overweight were 1.47 times (PHR = 1.47; 95% CI: 1.15 - 1.85) more at risk of dying compared to those who had normal weight. Children born to single mothers (PHR = 0.55; 95% CI: 0.24-0.99) and to mothers who used contraception (PHR = 0.51; 95% CI: 0.41 - 0.64) were less likely to die compared to their counterparts. The age of the mother at childbirth was marginally significant and a five-year increase in the mother's age reduced the chances of the child dying by 7% (PHR = 0.93; 95% CI: 0.85- 1.00). Being a Pentecostal (PHR = 0.61; 95% CI: 0.42-0.84) and being part of other religions not mentioned (PHR = 0.46; 95% CI: 0.24- 0.78) was protective from mortality.

With reference to the spatial adjusted Bayesian AFT model for the 2015 survey and adjusting for other factors, we found that multiple births (TR = 0.0019; 95% CI: 0.00036 - 0.011) was associated with shorter survival time of children and like in 2011, girls in 2015 still survived longer than boys (TR = 2.76; 95% CI: 1.07 - 6.25). Survival time of children born underweight (TR = 0.080; 95% CI: 0.024 - 0.24) and overweight (TR = 0.21; 95% CI: 0.077 - 0.54) was shorter than that of average weight children. Mother being single (TR = 10.23; 95% CI: 1.10 - 91.96), using contraceptives (TR = 11.84; 95% CI: 4.41 - 29.20), a five-year increase in mother's age (TR = 1.35; 95% CI: 1.10 - 1.78) increased the chances of a child surviving longer. Being affiliated to Pentecostalism (TR = 5.41; 95% CI: 1.48 - 18.86) and other religions (TR = 16.27; 95% CI: 2.59 - 105.84) was critical in longer survival times for children.

3.6 Spatial Distribution

3.6.1 2011 Spatial Distribution

The following spatial survival maps are for selected models from OpenBugs and R and they show U5M at the district level in Zimbabwe after adjusting for associated factors in 2011.

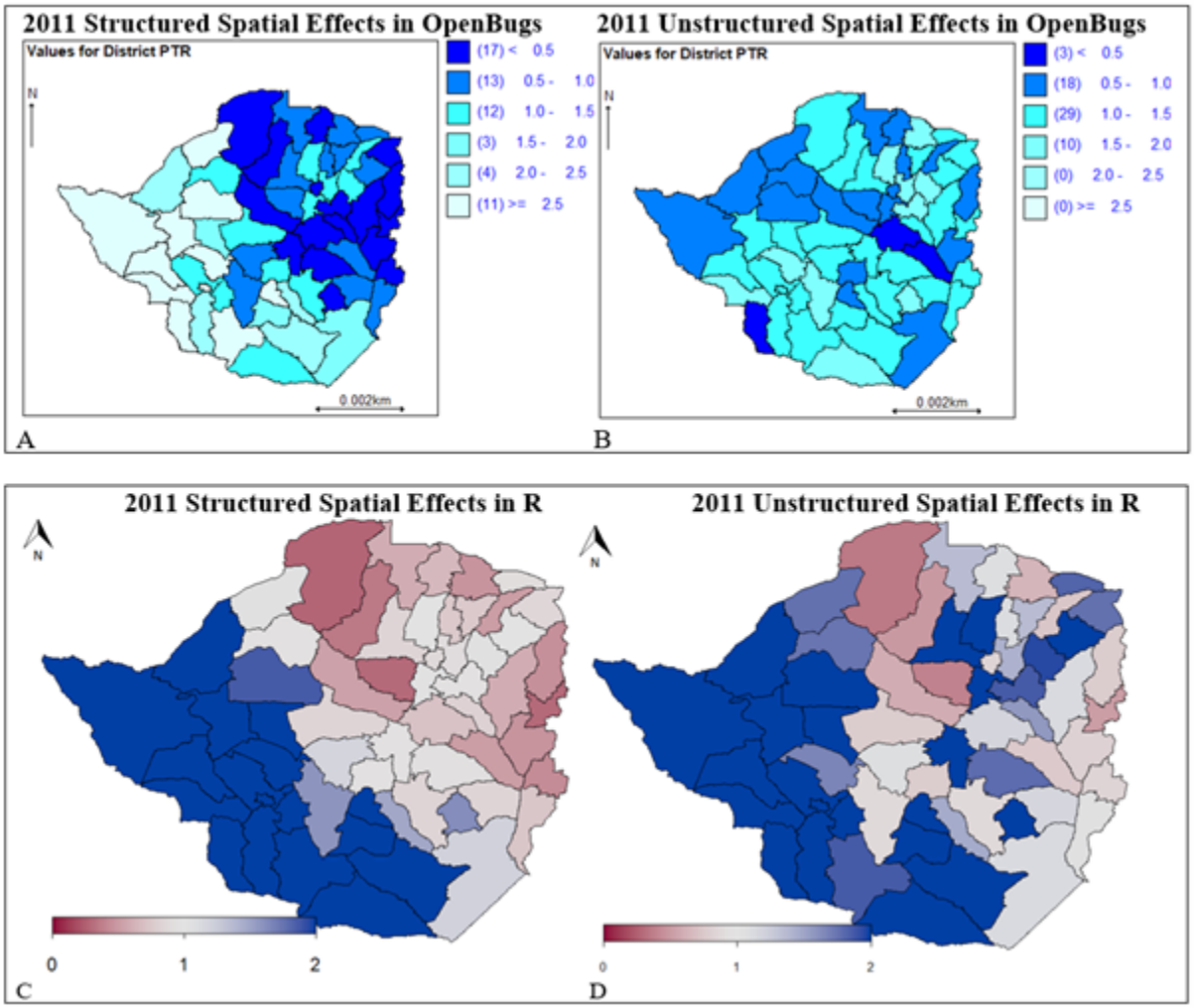


Figure 3.4: 2011 structured (left) and unstructured (right) spatial survival at district level from OpenBugs (top) and R (bottom).

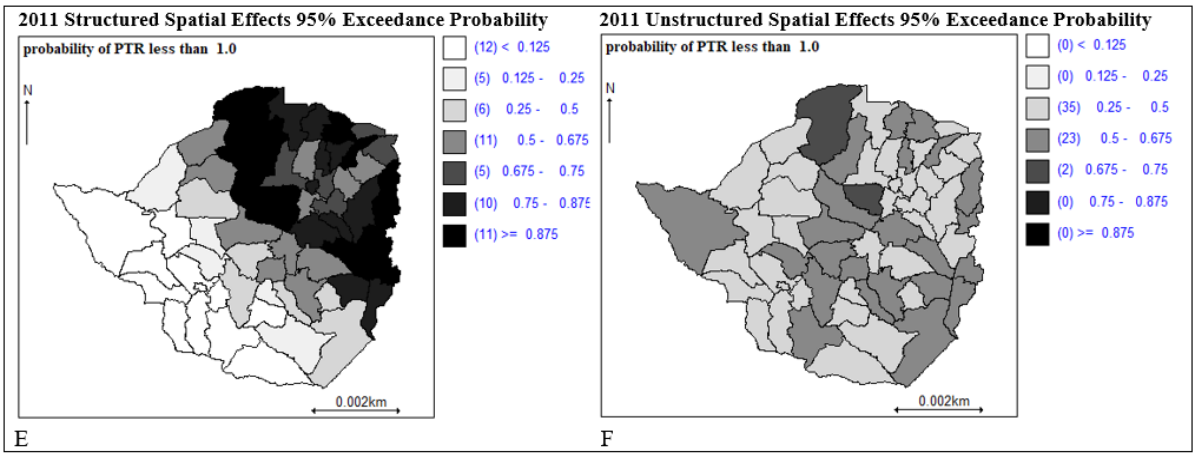


Figure 3.5: 2011 structured (left) and unstructured (right) effects maps of exceedance probability (PTR < 1.0) at district level.

The maps showing the distribution of district-level U5M in Zimbabwe for 2011 demographic and health surveys are shown in Figure 3.4 from both OpenBugs and R software. The deep blue colours denote hotspots of U5M for OpenBugs maps while for R maps, the red colours denote U5M hotspots. The maps of exceedence probabilities (PTR < 1) corresponding to the selected OpenBugs spatial models, are shown in Figure 3.5 with darker colours indicating higher probability of lower survival times. The posterior time ratios were used as a measure of effects. If we refer to maps A and E, the spatial distribution of U5M in 2011 exhibits lower survival times from the northern regions of the country in Mashonaland West and Mashonaland Central Provinces down through the central regions which include parts of Midlands Province and Harare Metropolitan moving down towards the eastern parts of the country mainly into Manicaland Province and south-eastern regions in Masvingo Province. The highest survival times were experienced in the western and southern regions of Zimbabwe, in districts from Bulawayo Metropolitan, Matabeleland North and Matabeleland South Provinces.

3.6.2 2015 Spatial Distribution

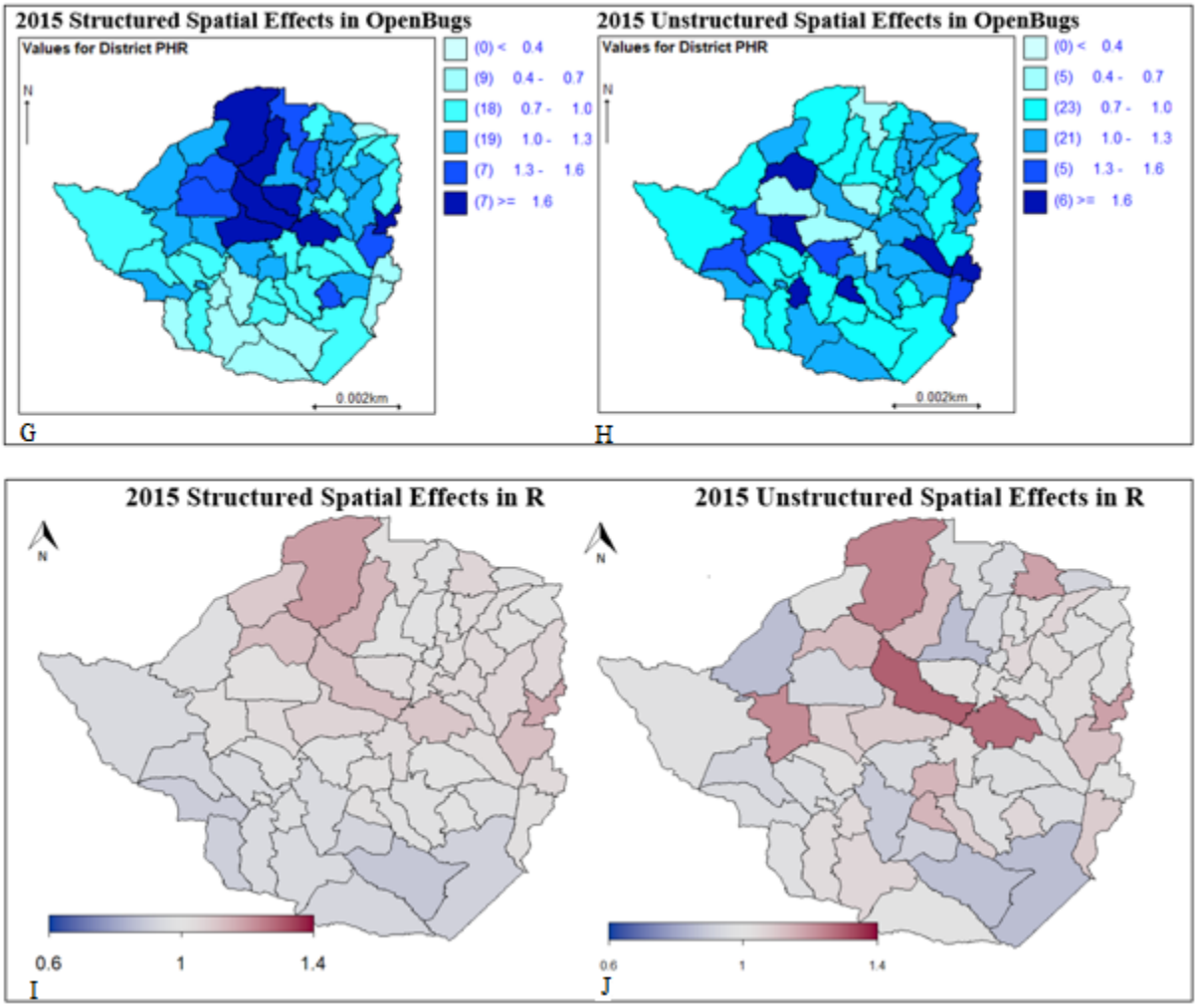


Figure 3.6: 2015 structured (left) and unstructured (right) spatial survival at district level from OpenBugs (top) and R (bottom).

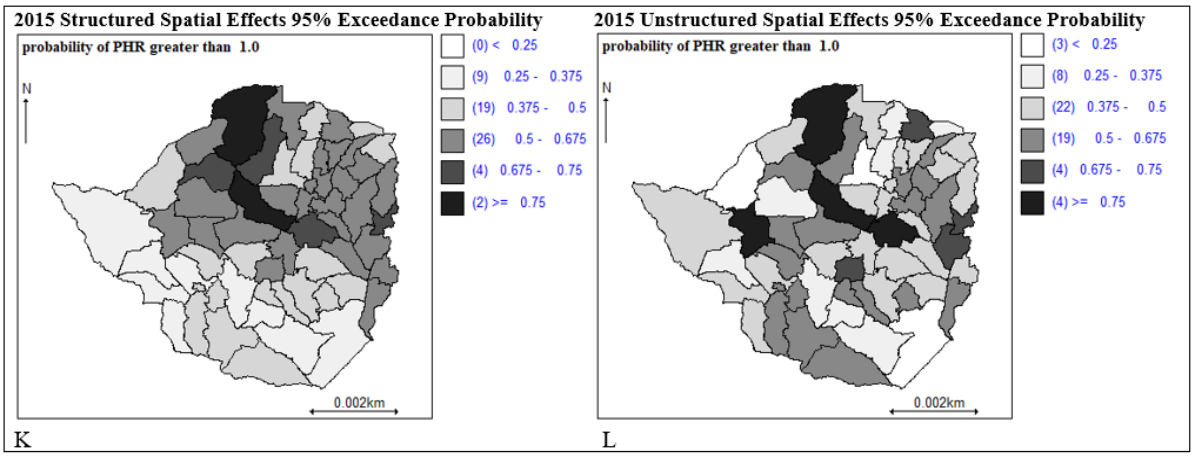


Figure 3.7: 2015 structured (left) and unstructured (right) effects maps of exceedance probability (PHR > 1.0) at district level.

The spatial survival maps given in Figure 3.6 show U5M mortality at the district level in Zimbabwe for selected adjusted 2015 models. The exceedence probabilities (PHR > 1) are shown in Figure 3.7 with darker colours showing higher probability of hazards. The distribution of child survival across the country's regions is shown in Figure 3.6 and hazard ratios were used as a measure of effect for the 2015 Cox-PH maps. Considering survival in 2015 from maps G and K, U5M was high in the northern (Mashonaland West Province), north eastern (Mashonaland Central), central (Midlands) and eastern (Manicaland Province) regions of the country and lower in the western (Matabeleland North) and southern (Matabeleland South) regions of the country. Child survival was worryingly low in Harare and Bulawayo Metropolitans in 2015.

Chapter 4

DISCUSSION and CONCLUSION

4.1 Introduction

In this section, we discuss key findings from our analysis of the two survey points and compare them to previous studies. We also discuss the strengths and limitations of the current study, and lastly, we conclude the study findings. The main aim of this study was to determine the geographical variations of U5M in Zimbabwe after adjusting for risk factors, comparing results from two survey points (2011 and 2015 ZDHS). Since the 2011 dataset failed the proportional hazards (PH) assumption test and the 2015 dataset satisfied the PH assumption, this chapter will be based on the AFT model from OpenBugs (model 6) for 2011 results and the Cox-PH model from OpenBugs (model 9) for 2015 results.

4.2 U5M Risk Factors

It is evident from the results of the current study that multiple births were a risk factor of U5M for both survey years and this is consistent with findings from previous studies in SSA which found a strong association between U5M and multiple births (41,71,72). Explanation to this may be based on that multiple births are classified as high-risk pregnancies and may be associated with complications such as preterm birth and babies from multiple births are at risk of being underweight, being born with defaults, being born disabled and all these factors are usually associated with mortality (29,72). Inspection of KM curves for the type of birth (Figure 3.4, panel A) showed rapid drops in the multiple birth sub-population during the early days of life which suggests that most deaths in multiple births occur early in life and might be due to complications at birth (73,74). As a result, mothers should be encouraged to utilize healthcare facilities and trained midwives, to reduce the adverse effects of birth complications.

Female children were at a lower risk of dying and this was comparable to past studies which found that survival was higher in females than in males (75–77). The explanation for this was centred around that girls are considered to be biologically stronger than boys. Due to their genetic makeup, girls were more resistant to infectious diseases and malnutrition, unlike boys who are more susceptible to diseases (73,75,76). The other explanation is that boys under-five years are more adventurous than girls and are eager to explore new things which makes them more vulnerable to accidental death (drowning, suffocating, traffic) (77,78).

For both surveys, the small birth size was a risk factor of U5M, and this was in line with findings from other studies (79,80). The effects of being small in size at birth could be viewed from the KM curves (Figure 3.4, panel C) and were more profound

during infancy (first year) than in childhood and according to some studies, this may be due to improved infant nutrition which helps small-sized babies to catch up with normal-sized babies in childhood (81). Mothers should therefore be educated on breastfeeding practices and should also be incorporated in feeding programs to balance the diet of both mothers and children to assist babies born underweight to gain substantially. However, conflicting results were obtained for large birth size as it had an insignificant effect on U5M in 2011 but was an important predictor of U5M in 2015. A study in Tanzania on fetal macrosomia found that mortality among large-sized babies at birth was significantly higher and they were more likely to experience neonatal complications such as birth trauma, asphyxia (brain damage due to lack of oxygen at birth) and respiratory distress (82).

The effect of mother's age at childbirth was negligible in 2011 but was substantial in 2015 where an increase in mother's age reduced the chances of U5M in children. Several studies in SSA have suggested that older mothers were more mature to take care of children (83–85). Young mothers should seek regular prenatal and postnatal checkups and family support for young mothers should be encouraged in childcare (86). A study conducted in low and middle-income countries suggested that older mothers were biologically mature to have children and they were more knowledgeable in childcare (85). Myrskyla et al. suggested that some women delay having children as they pursued education which made older mothers have higher education. The author further noted that older mothers had more access to social support and economic resources which aided child survival (87).

Unlike in other studies, where attainment of education by mother reduced the risks of U5M (9,12), in our study we could not find any significant differences in the effect of education on U5M which is however in line with studies done in low-income countries

(6) where the benefit of education is overshadowed by poverty. Contrary to other studies which have found marriage to be beneficiary in reducing U5M (7), in our study we found that U5M was lowest for single mothers. The results should, however, be used with caution as only a few single mothers participated in both surveys. Spousal support has been perceived to help in gathering of resources to improve child survival amongst couples. In a ten-year cohort study on the effect of marriage on infant mortality, it was found that single mothers were more prone to experiencing adverse outcomes when it came to child survival (88).

From our findings, contraceptive use by mothers was protective against U5M and children of mothers who used any form of contraception survived longer. This is consistent with several studies which indicated the importance of contraceptive use and family planning in increasing child survival (31,89–91). Studies on the relationship between contraceptive use, birth spacing and child survival in Bangladesh and India confirmed that contraceptive use increased birth spacing, and this was protective against U5M mortality (89,91). Mothers should therefore be educated on the benefits of family planning and contraceptive use as this reduces the chances of unplanned pregnancies and allows mothers to be better prepared to welcome a child into their lives. Some previous studies have also associated higher wealth status with higher child survival and lower U5M (92,93). However, the current study did not find any significant relationships between wealth index and U5M.

The current study found that in 2011, the rural residence was protective against the hazards of U5M but in 2015 the effect of residence was of negligible magnitude. Most studies have however found that the hazards were higher in rural areas and lower in urban areas as there was better access to resources including healthcare and sanitation in urban areas (94–96). Our results are opposed to what has been found

before and this may be ascribed to low progress in urban areas which cancels out the urban advantage (95), a situation that has been characteristic of Zimbabwe over the past years.

In this study, we found that the Apostolic sect had the highest U5M which is in line with several studies which found a strong association between Apostolic religion and high U5M (97,98). This could be attributed to the doctrines of the Apostolic religion which encouraged religious interventions for all people's problems (99). As a result, members are not eager to utilize health care services and facilities. They are reluctant to take their children for vaccinations, to seek healthcare for sick children, to use health facilities and personnel for child delivery, to get antenatal and postnatal care and to deal with child marriages, all of which could be used to explain the high U5M rates in the Apostolic sect (98–100). In light of this, it is of uttermost importance to engage religious leaders in health campaigns to reduce U5M and interventions should aim at breaking the faith and health barriers so there is a balance between the two.

4.3 Spatial Distribution of U5M

Geographical variations of U5M could be observed from the maps, with the northern regions, parts of middle regions and eastern regions being hotspots of U5M and the southern and western regions being coldspots for both survey points in Zimbabwe. The spatial distribution of U5M was similar for both years but hazards were much lower in 2015 than in 2011. Even though there is significant progress in reducing U5M across the country, districts in some provinces maintained high mortality rates and these were in the northern, north-eastern parts, central and eastern parts.

Negative progress was also observed in Harare Metropolitan which remained a hotspot

for both surveys, while Bulawayo Metropolitan remained an U5M coldspot in both surveys. Despite the negatives above, there was impressive progress in some districts in the eastern and south eastern regions, where the U5M hazards were cleared and some of 2011 hotspots became coldspots in 2015. There was notable progress in the southern and western parts of the country, where most districts remained U5M coldspots.

Zimbabwe is divided into farming regions and agro-ecological regions are the regions that receive high annual rainfall ($>$ 700mm) and are found from the central parts of the country, stretching to the northern, north-eastern, and eastern parts of the country, with the eastern highlands receiving the highest annual rainfall (101). Considering the geography of U5M from our study, we notice that these agro-ecological regions are hotspots of U5M. Evidence from the literature on U5M in SSA has shown an association between U5M and high rainfall (102). The reason might be that regions with high rainfall harbour a lot of diseases and standing waters are usually breeding ground for parasites. Tsiko analysed the spatial distribution of childhood diseases in Zimbabwe at the provincial level and found that in 2010-11, Manicaland Province was a hotspot for diarrheal diseases, cough and fever while Masvingo, Harare Metropolitan and Midlands Provinces were hotspots for Cough (26). In a study on spatial analysis of malaria cases in Zimbabwe from 2011 to 2016, Gwitira et al. reported that the northern, north-eastern, eastern and south-eastern districts of Zimbabwe were high-risk regions of Malaria and this had persisted over time (103). These studies on child morbidity in Zimbabwe may help explain the elevated U5M in some regions, since the areas which were identified as high-risk areas for child morbidity in Zimbabwe are the same areas that have been identified as hotspots for U5M in our study.

4.4 Comparison between 2011 and 2015 Findings

From the results of the two surveys, we found that the U5M rates were significantly lower in 2015 compared to 2011 implying that there was progress made in reducing U5M. One of the explanations to this might be that due to economic turmoil, Zimbabwe neglected the local currency in 2009 and adopted the multicurrency system, with the United States dollar (USD) as the main currency and this helped reduce inflation in the country and boosted economic growth (104). According to Tapera, the introduction of multicurrency system improved financial support and as a result of improved HIV/Aids financing, infant mortality dropped by up to 55% between 2009 and 2013 (105). The 2011 survey covered 2006-2010, a period during which the country was mainly using the Zimbabwean dollar and was faced with economic challenges which according to UN Zimbabwe made it difficult to put together human and medical resources to help fight U5M (106), while the 2015 survey covered 2011-2015, a period when the USD stabilised the economy meaning that by 2015 there were more funds for resource mobilization which aided in reducing U5M (105).

For U5M risk factors, type of birth, sex of child, small size of child at birth, mother's marital status, contraceptive use and religion were significant risk factors of U5M for both surveys while type of residence was significant in 2011 and not in 2015 and mother's age at childbirth and large size at birth were risk factors only in 2015. In terms of spatial distribution, there were similar hotspots and cold spots of U5M in 2011 and 2015 but the intensity of hazards was lower in 2015 than in 2011. Spatial clustering of U5M existed and children from districts in the northern, central and eastern regions of the country were at a high risk of experiencing hazards while staying in the western and southern regions was protective from hazards.

4.5 Modelling

Survival analysis techniques were applied in OpenBugs and R software to model U5M in Zimbabwe and the best fit models were selected. Bayesian estimation (BE) was used to fit our spatial survival models and the models performed better than the corresponding non-spatial models based on maximum likelihood estimation (MLE). It has been shown that if good selection of priors is done, then BE produces more consistent results than MLE (107). Cox-PH and AFT models from OpenBugs performed better and fit our data better than corresponding models in R software. The introduction of frailty allowed us to adjust for spatial random-effects (structured and unstructured) and the structured random-effects models performed better than the unstructured random-effects models. The unstructured spatial effects are assigned a Gaussian prior distribution under the assumption that the random effects are normally distributed, and this is not always the case since some random effects may violate the normality assumption and, in such situations, the structured effects with CAR prior may be more stable (108). We could not adjust for time-varying covariates (tvc) in OpenBugs, so instead of using Cox-PH with tvc where the PH-assumption was violated, we opted for AFT models as these have been used in past studies as an alternative to the Cox-PH models (44,109) in situations where the PH assumption is violated. Hence the use of both spatial Cox-PH and AFT models gave us more informative results.

4.6 Strengths and Limitations

4.6.1 Study Design

The data source is highly reliable, and the surveys had countrywide coverage with data collected by specially trained personnel from participants with high response rates. The retrospective nature of the study which covered a period of up to five years

prior to the survey meant that for some questions where official clinical records were not kept, the mothers had to rely on recall, and this might have led to recall bias. The current study focused on all-cause mortality as we could not infer causality for the cross-sectional survey data. For every DHS survey, there is a random selection of EAs and as a result, different EAs were used for both surveys as some of the 2011 EAs were not selected for the 2015 survey and this might have affected the comparability of the results from the two surveys. The high missingness for some important variables such as antenatal care and postnatal care led to the omission of the variables as we couldn't assume the data were missing at random, since there was a chance that mothers who lost their children did not bother answering that section.

4.6.2 Modelling Computational Aspects

The use of survival analysis incorporates the time survived by a child during the survey period and the contribution of the child is relative to the survival time unlike in logistic regression where the time component is not important. Spatial survival is an interesting area of study which is now getting more attention but has not been well studied like spatial linear and generalised linear models. Hence some of the functionality was not fully supported, so we could not adjust for sampling weights in the spatial models, and this might affect the accuracy of our results (110). OpenBugs and *spBayessurv* package in R did not provide the means for testing for the PH-assumptions for the spatial models, so we had to do the test in STATA. Even though OpenBugs had the best fit for our data, the computational times for OpenBugs models were very long when compared to R computational times which were shorter. The models also needed a high number of iterations for them to converge.

4.7 Conclusion

Information on the geographical distribution of U5M across the country is important in understanding how frailty due to unmeasured effects at the district level can affect child survival. The adjusted Bayesian spatial survival modeling provided better-fitting models than non-spatial models implying spatial clustering of U5M in Zimbabwe as mortality is similar for neighbouring districts. There were geographical variations in U5M across the country and we identified the northern, central and eastern parts to be hotspots that persisted between the two surveys. This points towards the need for implementation of strategies and targeted interventions at the subnational level to guide resource allocation to fast-track progress in reducing U5M.

Even though there was significant progress in decreasing U5M between 2011 and 2015, the mortality rates remained relatively high and if Zimbabwe is to achieve SDG 3 by 2030, then fundamental steps must be taken towards tackling the U5M risk factors. High-risk pregnancies (multiple births, young mothers) should be closely monitored, and mothers encouraged to seek pre-natal and post-natal care regularly. Women have to be engaged in contraceptive use and family planning programs to experience the benefits of these programs. The religion and healthcare barrier must be broken, and mothers should be educated on how to balance religion and healthcare to avoid adverse health outcomes.

References

1. Levels and Trends in Child Mortality Report 2018 United Nations Population Division Department of Economic and Social Affairs [Internet]. [cited 2019 Oct 25]. Available from: https://www.un.org/en/development/desa/population/publications/mortality/child-mortality-report-2018.asp

2. Molyneux M, Molyneux E. Reaching Millennium Development Goal 4. The Lancet Global Health [Internet]. 2016 Mar [cited 2020 Aug 16];4(3):e146–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2214109X16000097

3. Child survival and the SDGs [Internet]. UNICEF DATA. [cited 2021 Jan 12]. Available from: https://data.unicef.org/topic/child-survival/child-survival-sdgs/

4. Haley CA, Vermund SH, Moyo P, Kipp AM, Madzima B, Kanyowa T, et al. Impact of a critical health workforce shortage on child health in Zimbabwe: A country case study on progress in child survival, 2000–2013. Health Policy and Planning [Internet]. 2017 Jan [cited 2020 Aug 16];czw162. Available from: https://academic.oup.com/heapol/article-lookup/doi/10.1093/heapol/czw162

5. O'Hare B, Makuta I. An analysis of the potential for achieving the fourth millennium development goal in SSA with domestic resources. Globalization and Health [Internet]. 2015 [cited 2020 Aug 16];11(1):8. Available from: http:

//www.globalizationandhealth.com/content/11/1/8

6. Kandala NB, Mandungu T.P., Mbela K., Nzita K.P., Kalambayi B.B., Kayembe K.P. Child mortality in the Democratic Republic of Congo: Cross-sectional evidence of the effect of geographic location and prolonged conflict from a national household survey. BMC public health [Internet]. 2014;14:266. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed12&NEWS=N&AN=24649944

7. Kanmiki EW, Bawah AA, Agorinya I, Achana FS, Awoonor-williams JK, Oduro AR, et al. Socio-economic and demographic determinants of under-five mortality in rural northern Ghana. BMC International Health and Human Rights [Internet]. 2014 Aug [cited 2021 Aug 4];14(1):24. Available from: https://doi.org/10.1186/1472-698X-14-24

8. Anyamele OD, Ukawuilulu JO, Akanegbu B. The Role of Wealth and Mother's Education in Infant and Child Mortality in 26 Sub-Saharan African Countries: Evidence from Pooled Demographic and Health Survey (DHS) Data 2003–2011 and African Development Indicators (ADI), 2012. Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement [Internet]. 2017 [cited 2021 Aug 5];130(3):1125–46. Available from: https://econpapers.repec.org/article/sprsoinre/v_3a130_3ay_3a2017_3ai_3a3_3ad_3a10.1007_5fs11205-015-1225-x.htm

9. Wegbom A, Essi I, Kiri V. Survival Analysis of Under-five Mortality and Its Associated Determinants in Nigeria: Evidence from a Survey Data. 2019 May;9(2):59–66.

10. Ytterstad E. Frailty in Survival Analysis of Widowhood Mortality. Journal of Probability and Statistics [Internet]. 2018 [cited 2021 Aug 5];2018:1–5. Available from: https://www.mendeley.com/catalogue/f9b359a8-3da3-39ec-

9b03-a4188607fb7c/

11. Zimbabwe National statistics agency and IDF. Zimbabwe Demographics and Health Survey 2015: Key Indicators. 2016; Available from: https://zimbabwe.unfpa.org/sites/default/files/pub-pdf/ZDHS%20Preliminary%20Results.pdf

12. Grépin KA, Bharadwaj P. Maternal education and child mortality in Zimbabwe. Journal of Health Economics. 2015 Dec;44:97–117.

13. Makate M, Makate C. The impact of prenatal care quality on neonatal, infant and child mortality in Zimbabwe: Evidence from the demographic and health surveys. Health Policy and Planning. 2017 Apr;32(3):395–404.

14. Nasejje JB, Mwambi HG, Achia TNO. Understanding the determinants of under-five child mortality in Uganda including the estimation of unobserved household and community effects using both frequentist and Bayesian survival analysis approaches. BMC Public Health [Internet]. 2015 Oct [cited 2019 Oct 29];15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4591593/

15. Kembo J, Van Ginneken JK. Determinants of infant and child mortality in Zimbabwe: Results of multivariate hazard analysis. Demographic Research. 2009;21(July 2009):367–84.

16. Mosley WH, Chen LC. An Analytical Framework for the Study of Child Survival in Developing Countries. Population and Development Review. 1984;10:25–45.

17. Lamichhane R, Zhao Y, Paudel S, Adewuyi EO. Factors associated with infant mortality in Nepal: A comparative analysis of Nepal demographic and health surveys (NDHS) 2006 and 2011 [Internet]. Vol. 17, BMC Public Health. 2017 [cited 2021 Aug 5]. p. 53. Available from: https://doi.org/10.1186/s12889-016-3922-z

18. Li Z, Hsiao Y, Godwin J, Martin BD, Wakefield J, Clark SJ, et al. Changes

in the spatial distribution of the under-five mortality rate: Small-area analysis of 122 DHS surveys in 262 subregions of 35 countries in Africa. Uthman O, editor. PLOS ONE [Internet]. 2019 Jan [cited 2019 Oct 24];14(1):e0210645. Available from: http://dx.plos.org/10.1371/journal.pone.0210645

19. Yaya S, Bishwajit G, Okonofua F, Uthman OA. Under five mortality patterns and associated maternal risk factors in sub-Saharan Africa: A multi-country analysis. Wouwe JP van, editor. PLOS ONE [Internet]. 2018 Oct [cited 2021 Jan 8];13(10):e0205977. Available from: https://dx.plos.org/10.1371/journal.pone.0205977

20. Ogbo FA, Ezeh OK, Awosemo AO, Ifegwu IK, Tan L, Jessa E, et al. Determinants of trends in neonatal, post-neonatal, infant, child and under-five mortalities in Tanzania from 2004 to 2016. BMC Public Health [Internet]. 2019 Dec [cited 2021 Jan 8];19(1):1243. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-7547-x

21. Nigatu D, Haile D, Gebremichael B, M Tiruneh Y. Predictive accuracy of perceived baby birth size for birth weight: A cross-sectional study from the 2016 Ethiopian Demographic and Health Survey. BMJ Open [Internet]. 2019 Dec [cited 2021 Jan 8];9(12):e031986. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2019-031986

22. Oral E, Cağdaş A, Gezer A, Kaleli S, Aydinli K, Öçer F. Perinatal and maternal outcomes of fetal macrosomia. European Journal of Obstetrics & Gynecology and Reproductive Biology [Internet]. 2001 Dec [cited 2021 Feb 5];99(2):167–71. Available from: https://linkinghub.elsevier.com/retrieve/pii/S030121150100416X

23. Zhang X, Decker A, Platt RW, Kramer MS. How big is too big? The perinatal consequences of fetal macrosomia. American Journal of Obstetrics and Gynecology

[Internet]. 2008 May [cited 2021 Feb 5];198(5):517.e1–6. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002937807022703

24. Kazembe LN, Kandala NB. Estimating areas of common risk in low birth weight and infant mortality in Namibia: A joint spatial analysis at sub-regional level. Spatial and Spatio-temporal Epidemiology [Internet]. 2015;12:27–37. Available from: http://dx.doi.org/10.1016/j.sste.2015.02.001

25. Azuine R. Exclusive Breastfeeding and Under-Five Mortality, 2006-2014: A Cross-National Analysis of 57 Low- and-Middle Income Countries. International Journal of MCH and AIDS (IJMA) [Internet]. 2015 [cited 2021 Jan 8];4(1). Available from: http://mchandaids.org/index.php/IJMA/article/view/52

26. Tsiko RG. Bayesian spatial analysis of childhood diseases in Zimbabwe Biostatistics and methods. BMC Public Health [Internet]. 2015;15(1). Available from: http://dx.doi.org/10.1186/s12889-015-2182-7

27. Musenge E, Vounatsou P, Collinson M, Tollman S, Kahn K. The contribution of spatial analysis to understanding HIV/TB mortality in children: A structural equation modelling approach. Global Health Action [Internet]. 2013 Jan [cited 2021 Aug 5];6(19266):10.3402/gha.v6i0.19266. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3556702/

28. Dwomoh D, Amuasi S, Agyabeng K, Incoom G, Alhassan Y, Yawson AE. Understanding the determinants of infant and under-five mortality rates: A multivariate decomposition analysis of Demographic and Health Surveys in Ghana, 2003, 2008 and 2014. BMJ global health. 2019;4(4):e001658.

29. Hong R. Effect of multiple birth on infant mortality in Bangladesh. Journal of paediatrics and child health. 2006 Nov;42:630–5.

30. Arku RE, Bennett JE, Castro MC, Agyeman-Duah K, Mintah SE, Ware JH,

et al. Geographical Inequalities and Social and Environmental Risk Factors for Under-Five Mortality in Ghana in 2000 and 2010: Bayesian Spatial Analysis of Census Data. PLoS Medicine. 2016;13(6):1–14.

31. Chikandiwa A, Burgess E, Otwombe K, Chimoyi L. Use of contraceptives, high risk births and under-five mortality in Sub Saharan Africa: Evidence from Kenyan (2014) and Zimbabwean (2011) demographic health surveys. BMC women's health. 2018 Oct;18(1):173.

32. Gayawan E, Adarabioyo MI, Okewole DM, Fashoto SG, Ukaegbu JC. Geographical variations in infant and child mortality in West Africa: A geo-additive discrete-time survival modelling. Genus [Internet]. 2016;72(1). Available from: http://dx.doi.org/10.1186/s41118-016-0009-8

33. Zimbabwe - Multiple Indicator Cluster Survey 2014. [cited 2021 Sep 2]; Available from: https://microdata.worldbank.org/index.php/catalog/2527

34. Omariba DWR, Beaujot R, Rajulton F. Determinants of Infant and Child Mortality in Kenya: An Analysis Controlling for Frailty Effects. Population Research and Policy Review [Internet]. 2007 [cited 2021 Aug 5];26(3):299–321. Available from: https://www.jstor.org/stable/40230978

35. Hougaard P. Frailty models for survival data. Lifetime Data Analysis. 1995;1(3):255–73.

36. Zhou H, Hanson T, Zhang J. Generalized accelerated failure time spatial frailty model for arbitrarily censored data. Lifetime Data Analysis. 2017;23(3):495–515.

37. Ayele D, Zewotir T, Mwambi H. Survival analysis of under-five mortality using Cox and frailty models in Ethiopia. Journal of Health, Population, and Nutrition. 2017 Jun;36(1):25.

38. Banerjee S, Wall MM, Carlin BP. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. Biostatistics [Internet]. 2003 Jan [cited 2021 Aug 5];4(1):123–42. Available from: https://doi.org/10.1093/biostatistics/4.1.123

39. Kneib T. Mixed model based inference in structured additive regression [Internet] [Text.PhD thesis]. Ludwig-Maximilians-Universität München; 2006 [cited 2021 Aug 5]. Available from: https://edoc.ub.uni-muenchen.de/5011/

40. Kandala N-B, Ghilagaber G. A Geo-Additive Bayesian Discrete-Time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi. Quality & Quantity [Internet]. 2006 Dec [cited 2021 Jan 12];40(6):935–57. Available from: http://link.springer.com/10.1007/s11135-005-3268-6

41. Kazembe L, Clarke A, Kandala N-B. Childhood mortality in sub-Saharan Africa: Cross-sectional insight into small-scale geographical inequalities from Census data. BMJ Open [Internet]. 2012 [cited 2021 Jan 12];2(5):e001421. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2012-001421

42. Chen. A Bayesian Approach for the Cox Proportional Hazards Model with Covariates Subject to Detection Limit. International Journal of Statistics in Medical Research [Internet]. 2014 [cited 2020 Jul 11]; Available from: http://www.lifescienceglobal.com/journals/international-journal-of-statistics-in-medical-research/volume-3-number-1/91-abstract/ijsmr/658-abstract-a-bayesian-approach-for-the-cox-proportional-hazards-model-with-covariates-subject-to-detection-limit

43. Mostafa A. Fitting The Hazard Ratio for Cluster Survival Data with Frailty Effect Via WinBUGS. J Basic Appl Sci Res. 2012 Jan;2:5781–93.

44. Zhang J, Lawson AB. Bayesian parametric accelerated failure time spatial model

and its application to prostate cancer. Journal of Applied Statistics [Internet]. 2011 Mar [cited 2020 Jul 11];38(3):591–603. Available from: http://www.tandfonline.com/doi/abs/10.1080/02664760903521476

45. Prabhash K, M Patil V, Noronha V, Joshi A, Bhattacharjee A. Bayesian Accelerated Failure Time and Its Application in Chemotherapy Drug Treatment Trial. Statistics in Transition New Series [Internet]. 2016 [cited 2020 Jul 11];17(4):671–90. Available from: https://www.exeley.com/statistics_in_transition/doi/10.21307/stattrans-2016-046

46. Zhang J, Lawson AB. Bayesian parametric accelerated failure time spatial model and its application to prostate cancer. Journal of Applied Statistics [Internet]. 2011 Mar [cited 2020 Oct 31];38(3):591–603. Available from: http://www.tandfonline.com/doi/abs/10.1080/02664760903521476

47. Zhou H, Hanson T, Zhang J. spBayesSurv: Fitting Bayesian Spatial Survival Models Using R. Journal of Statistical Software [Internet]. 2020 Feb [cited 2021 Aug 5];92(1):1–33. Available from: https://www.jstatsoft.org/index.php/jss/article/view/v092i09

48. Zimstat ZNSA, International ICF. Zimbabwe Demographic and Health Survey 2010-11. 2012 Mar [cited 2019 Oct 29]; Available from: https://dhsprogram.com/publications/publication-FR254-DHS-Final-Reports.cfm

49. Corsi DJ, Neuman M, Finlay JE, Subramanian S. Demographic and health surveys: A profile. International Journal of Epidemiology [Internet]. 2012 Dec [cited 2020 Aug 15];41(6):1602–13. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys184

50. Buuren S van, Groothuis-Oudshoorn K. **Mice** : Multivariate Imputation by Chained Equations in *r*. Journal of Statistical Software [Internet]. 2011 [cited 2020

Jul 12];45(3). Available from: http://www.jstatsoft.org/v45/i03/

51. DIVA-GIS free, simple & effective [Internet]. [cited 2020 Nov 12]. Available from: https://www.diva-gis.org/

52. The DHS Program - Quality information to plan, monitor and improve population, health, and nutrition programs [Internet]. [cited 2020 Nov 12]. Available from: https://dhsprogram.com/

53. Welcome to the QGIS project! [Internet]. [cited 2020 Nov 12]. Available from: https://qgis.org/en/site/

54. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.R-project.org/

55. Autoplot.survfit: Autoplot 'survival::Survfit' in ggfortify: Data Visualization Tools for Statistical Analysis Results [Internet]. [cited 2020 Oct 31]. Available from: https://rdrr.io/cran/ggfortify/man/autoplot.survfit.html

56. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing [Internet]. 2000 Oct [cited 2020 Nov 12];10(4):325–37. Available from: https://doi.org/10.1023/A:1008929526011

57. Etikan I. The Kaplan Meier Estimate in Survival Analysis. Biometrics & Biostatistics International Journal [Internet]. 2017 Feb [cited 2020 Jul 11];5(2). Available from: https://medcraveonline.com/BBIJ/the-kaplan-meier-estimate-in-survival-analysis.html

58. Rich JT, Neely JG, Paniello RC, Voelker CCJ, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngology–Head and

Neck Surgery [Internet]. 2010 Sep [cited 2020 Jul 11];143(3):331–6. Available from: http://journals.sagepub.com/doi/10.1016/j.otohns.2010.05.007

59. Stel VS, Dekker FW, Tripepi G, Zoccali C, Jager KJ. Survival Analysis I: The Kaplan-Meier Method. Nephron Clinical Practice [Internet]. 2011 [cited 2020 Jul 11];119(1):c83–8. Available from: https://www.karger.com/Article/FullText/324758

60. Cox DR. Regression Models and Life-Tables. Journal of the Royal Statistical Society Series B (Methodological) [Internet]. 1972 [cited 2020 Nov 12];34(2):187–220. Available from: https://www.jstor.org/stable/2985181

61. Besag J, Kooperberg C. On Conditional and Intrinsic Autoregression. Biometrika [Internet]. 1995 Dec [cited 2021 Apr 12];82(4):733. Available from: https://www.jstor.org/stable/2337341?origin=crossref

62. Morris M, Wheeler-Martin K, Simpson D, Mooney SJ, Gelman A, DiMaggio C. Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. Spatial and Spatio-temporal Epidemiology [Internet]. 2019 Nov [cited 2021 Apr 12];31:100301. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1877584518301175

63. Lavine ML, Hodges JS. On Rigorous Specification of ICAR Models. The American Statistician [Internet]. 2012 Feb [cited 2021 Apr 12];66(1):42–9. Available from: http://www.tandfonline.com/doi/abs/10.1080/00031305.2012.654746

64. Spatial Models in Stan: Intrinsic Auto-Regressive Models for Areal Data [Internet]. [cited 2021 Apr 12]. Available from: https://mc-stan.org/users/documentation/case-studies/icar_stan.html

65. Kassa TH. Bayesian Survival Analysis of Diabetes Mellitus Patients: A Case Study in Tikur Anbessa Specialized Hospital, Addis Ababa, Ethiopia. Journal of

Reliability and Statistical Studies [Internet]. 2018 Dec [cited 2021 Aug 5];37–56. Available from: https://www.journal.riverpublishers.com/index.php/JRSS

66. Alvares D, Lázaro E, Gómez-Rubio V, Armero C. Bayesian survival analysis with BUGS. arXiv:200505952 [stat] [Internet]. 2020 May [cited 2020 Jul 11]; Available from: http://arxiv.org/abs/2005.05952

67. Banbeta A, Seyoum D, Belachew T, Birlie B, Getachew Y. Modeling time-to-cure from severe acute malnutrition: Application of various parametric frailty models. Archives of Public Health [Internet]. 2015 Dec [cited 2020 Jul 11];73(1):6. Available from: https://archpublichealth.biomedcentral.com/articles/10.1186/2049-3258-73-6

68. Wang Y. Estimation of accelerated failure time models with random effects [Internet] [Doctor of Philosophy]. [Ames]: Iowa State University, Digital Repository; 2006 [cited 2020 Jul 11]. Available from: https://lib.dr.iastate.edu/rtd/3062/

69. Mostafa A, Ben Ghorbal A. Using WinBUGS to Cox Model with Changing from the Baseline Hazard Function. Applied Mathematical Sciences. 2011 Jan;5:2217–40.

70. Abate L, Tadesse M. Application of Bayesian Approach Survival Analysis of Under-five Pneumonia Patients in Tercha General Hospital, South West Ethiopia. International Journal on Data Science and Technology [Internet]. 2020 [cited 2020 Aug 16];6(1):44. Available from: http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=390&doi=10.11648/j.ijdst.20200601.16

71. Uthman OA, Uthman MB, Yahaya I. A population-based study of effect of multiple birth on infant mortality in Nigeria. BMC Pregnancy and Childbirth [Internet]. 2008 Dec [cited 2021 Feb 4];8(1):41. Available from: http://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/1471-2393-8-41

72. Dwomoh D, Amuasi S, Agyabeng K, Incoom G, Alhassan Y, Yawson AE. Understanding the determinants of infant and under-five mortality rates: A multivariate decomposition analysis of Demographic and Health Surveys in Ghana, 2003, 2008 and 2014. BMJ Global Health [Internet]. 2019 Jul [cited 2021 Feb 4];4(4):e001658. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2019-001658

73. Pongou R. Why Is Infant Mortality Higher in Boys Than in Girls? A New Hypothesis Based on Preconception Environment and Evidence From a Large Sample of Twins. Demography [Internet]. 2013 Apr [cited 2021 Feb 4];50(2):421–44. Available from: http://link.springer.com/10.1007/s13524-012-0161-5

74. Smith LK, Manktelow BN, Draper ES, Boyle EM, Johnson SJ, Field DJ. Trends in the incidence and mortality of multiple births by socioeconomic deprivation and maternal age in England: Population-based cohort study. BMJ Open [Internet]. 2014 Apr [cited 2021 Feb 4];4(4):e004514. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2013-004514

75. Sex Differentials in Childhood Mortality United Nations Population Division Department of Economic and Social Affairs [Internet]. [cited 2021 Feb 4]. Available from: https://www.un.org/en/development/desa/population/publications/mortality/sexdifferentials.asp

76. Costa JC, Silva ICM da, Victora CG. Gender bias in under-five mortality in low/middle-income countries. BMJ Global Health [Internet]. 2017 Jul [cited 2021 Feb 4];2(2):e000350. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2017-000350

77. Lili X, Jian H, Liping L, Zhiyu L, Hua W. Epidemiology of Injury-Related Death in Children under 5 Years of Age in Hunan Province, China, 2009–2014. Simeoni

U, editor. PLOS ONE [Internet]. 2017 Jan [cited 2021 Feb 4];12(1):e0168524. Available from: https://dx.plos.org/10.1371/journal.pone.0168524

78. Al-Thani M, Al-Thani A-A, Toumi A, Khalifa S, Ijaz MA, Akram H. Trends in the leading causes of childhood mortality from 2004 to 2016 in Qatar. Central Asian Journal of Global Health. 2018;7(1):334.

79. Ezeh OK, Agho KE, Dibley MJ, Hall JJ, Page AN. Risk factors for postneonatal, infant, child and under-5 mortality in Nigeria: A pooled cross-sectional analysis. BMJ Open [Internet]. 2015 Mar [cited 2021 Feb 5];5(3):e006779–9. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2014-006779

80. Khan FA, Mullany LC, Wu LF-S, Ali H, Shaikh S, Alland K, et al. Predictors of neonatal mortality: Development and validation of prognostic models using prospective data from rural Bangladesh. BMJ Global Health [Internet]. 2020 Jan [cited 2021 Feb 5];5(1):e001983. Available from: https://gh.bmj.com/lookup/doi/10.1136/bmjgh-2019-001983

81. Wilcox AJ. On the importance—and the unimportance— of birthweight. International Journal of Epidemiology [Internet]. 2001 Dec [cited 2021 Feb 5];30(6):1233–41. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/30.6.1233

82. Said AS, Manji KP. Risk factors and outcomes of fetal macrosomia in a tertiary centre in Tanzania: A case-control study. BMC Pregnancy and Childbirth [Internet]. 2016 Aug [cited 2021 Aug 5];16(1):243. Available from: https://doi.org/10.1186/s12884-016-1044-3

83. Ayotunde T, Mary O, Melvin A, Faniyi F. Maternal Age At Birth And Under-5 Mortality In Nigeria. East African Journal of Public Health [Internet]. 2009 Sep [cited 2021 Feb 12];6(1). Available from: http://www.ajol.info/index.php/

eajph/article/view/45735

84. Sinha S, Aggarwal AR, Osmond C, Fall CHD, Bhargava SK, Sachdev HS. Maternal age at childbirth and perinatal and under-five mortality in a prospective birth cohort from Delhi. Indian Pediatrics [Internet]. 2016 Oct [cited 2021 Feb 12];53(10):871–7. Available from: http://link.springer.com/10.1007/s13312-016-0950-9

85. Finlay JE, Ozaltin E, Canning D. The association of maternal age with infant mortality, child anthropometric failure, diarrhoea and anaemia for first births: Evidence from 55 low- and middle-income countries. BMJ Open [Internet]. 2011 Sep [cited 2021 Feb 12];1(2):e000226–6. Available from: https://bmjopen.bmj.com/lookup/doi/10.1136/bmjopen-2011-000226

86. Angley M, Divney A, Magriples U, Kershaw T. Social Support, Family Functioning and Parenting Competence in Adolescent Parents. Maternal and Child Health Journal [Internet]. 2015 Jan [cited 2021 Apr 23];19(1):67–73. Available from: http://link.springer.com/10.1007/s10995-014-1496-x

87. Myrskylä M, Barclay K, Goisis A. Advantages of later motherhood. Der Gynäkologe [Internet]. 2017 Oct [cited 2021 Feb 12];50(10):767–72. Available from: http://link.springer.com/10.1007/s00129-017-4124-1

88. Balayla J, Azoulay L, Abenhaim HA. Maternal Marital Status and the Risk of Stillbirth and Infant Death: A Population-Based Cohort Study on 40 Million Births in the United States. Women's Health Issues [Internet]. 2011 Sep [cited 2021 Feb 9];21(5):361–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1049386711000855

89. Shukla A, Kumar A, Mozumdar A, Aruldas K, Acharya R, Ram F, et al. Association between modern contraceptive use and child mortality in India: A calendar data analysis of the National Family Health Survey (2015-16). SSM - Popula-

tion Health [Internet]. 2020 Aug [cited 2021 Feb 9];11:100588. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2352827319302265

90. Bitew F, Nyarko SH. Modern contraceptive use and intention to use: Implication for under-five mortality in Ethiopia. Heliyon [Internet]. 2019 Aug [cited 2021 Feb 9];5(8):e02295. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2405844019359559

91. Saha UR, Soest A van. Contraceptive Use, Birth Spacing, and Child Survival in Matlab, Bangladesh. Studies in Family Planning [Internet]. 2013 Mar [cited 2021 Feb 9];44(1):45–66. Available from: http://doi.wiley.com/10.1111/j.1728-4465.2013.00343.x

92. Ekholuenetale M, Wegbom AI, Tudeme G, Onikan A. Household factors associated with infant and under-five mortality in sub-Saharan Africa countries. International Journal of Child Care and Education Policy [Internet]. 2020 Dec [cited 2021 Mar 12];14(1):10. Available from: https://ijccep.springeropen.com/articles/10.1186/s40723-020-00075-1

93. Bado AR, Appunni SS. Decomposing Wealth-Based Inequalities in Under-Five Mortality in West Africa. Iranian Journal of Public Health. 2015 Jul;44(7):920–30.

94. Dendup T, Zhao Y, Putra IGNE. Rural-urban differentials in the determinants of under-five mortality in Bhutan. Journal of Health Research [Internet]. 2021 Sep [cited 2021 Feb 16];35(3):226–39. Available from: https://doi.org/10.1108/JHR-09-2019-0208

95. Forde I, Tripathi V. Association of Place of Residence and Under-Five Mortality in Middle- and Low-Income Countries: A Meta-Analysis. Children [Internet]. 2018 Apr [cited 2021 Feb 16];5(4):51. Available from: http://www.mdpi.com/2227-9067/5/4/51

96. Ettarh RR, Kimani J. Determinants of under-five mortality in rural and urban Kenya. Rural and Remote Health. 2012;12:1812.

97. Ha W, Salama P, Gwavuya S, Kanjala C. Is religion the forgotten variable in maternal and child health? Evidence from Zimbabwe. Social Science & Medicine [Internet]. 2014 Oct [cited 2021 Feb 10];118:80–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0277953614005061

98. Ha W, Gwavuya S, Salama P. Can religion kill? The association between membership of the Apostolic faith and child mortality in Zimbabwe. Journal of Public Health in Africa [Internet]. 2018 Oct [cited 2021 Aug 5];9(2):707. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6325421/

99. Hallfors DD, Iritani BJ, Zhang L, Hartman S, Luseno WK, Mpofu E, et al. "I thought if I marry the prophet I would not die": The significance of religious affiliation on marriage, HIV testing, and reproductive health practices among young married women in Zimbabwe. SAHARA-J: Journal of Social Aspects of HIV/AIDS [Internet]. 2016 Jan [cited 2021 Mar 5];13(1):178–87. Available from: https://www.tandfonline.com/doi/full/10.1080/17290376.2016.1245627

100. Munyaradzi Kenneth D, Marvellous M, Stanzia M, Memory D-M. Praying until Death: Apostolicism, Delays and Maternal Mortality in Zimbabwe. PLoS ONE [Internet]. 2016 Aug [cited 2021 Mar 5];11(8). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4979998/

101. Kashangura C. POLICY ISSUES FOR POSSIBLE TEMPORAL AND SPATIAL RAINFALL DISTRIBUTION CHANGES DUE TO CLIMATE CHANGE. 2014 [cited 2021 Feb 19]; Available from: http://rgdoi.net/10.13140/2.1.2701.4724

102. Egondi T, Kyobutungi C, Kovats S, Muindi K, Ettarh R, Rocklöv J. Time-series analysis of weather and mortality patterns in Nairobi's informal settlements. Global

Health Action [Internet]. 2012 Dec [cited 2021 Mar 12];5(1):19065. Available from: https://www.tandfonline.com/doi/full/10.3402/gha.v5i0.19065

103. Gwitira I, Mukonoweshuro M, Mapako G, Shekede MD, Chirenda J, Mberikunashe J. Spatial and spatio-temporal analysis of malaria cases in Zimbabwe. Infectious Diseases of Poverty [Internet]. 2020 Dec [cited 2021 Feb 19];9(1):146. Available from: https://idpjournal.biomedcentral.com/articles/10.1186/s40249-020-00764-6

104. Mpofu RT. Dollarization and economic development in Zimbabwe: An interrupted time-series analysis. Risk Governance and Control: Financial Markets and Institutions [Internet]. 2015 [cited 2021 Mar 5];5(4):38–48. Available from: http://www.virtusinterpress.org/DOLLARIZATION-AND-ECONOMIC.html

105. Tapera O. Insensitivity of Infant Mortality Rates to Improved HIV/Aids Financing: Evidence from Post-Dollarisation Zimbabwe 2009-2013. International Journal of Developing Societies. 2014;3(2):22–8.

106. Millennium Development Goal 4 [Internet]. UNDP in Zimbabwe. [cited 2021 Apr 15]. Available from: https://www.zw.undp.org/content/zimbabwe/en/home/post-2015/mdgoverview/overview/mdg4.html

107. Ramadhini F, Djuraidah A, Wigena AH. Simulation Studyfor Comparison of Maximum Likelihood and Bayesian Method in Spatial Autoregressive Models with Heteroskedasticity. In: Proceedings of the Proceedings of the 1st International Conference on Statistics and Analytics, ICSA 2019, 2-3 August 2019, Bogor, Indonesia [Internet]. Bogor, Indonesia: EAI; 2020 [cited 2021 Apr 23]. Available from: http://eudl.eu/doi/10.4108/eai.2-8-2019.2290484

108. Cheruiyot TB, Samuel M, Anthony W. A More Robust Random Effects Model for Disease Mapping. American Journal of Theoretical and Applied

Statistics [Internet]. 2018 Jan [cited 2021 Aug 6];7(1):29. Available from: http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=146&doi=10.11648/j.ajtas.20180701.14

109. Fisher G, Lawson AB. Bayesian modeling of georeferenced cancer survival. Annals of Cancer Epidemiology [Internet]. 2020 Jun [cited 2021 Mar 4];4:6–6. Available from: http://ace.amegroups.com/article/view/5697/html

110. Bell BA, Onwuegbuzie AJ, Ferron JM, Jiao QG, Hibbard ST, Kromrey JD. Use of Design Effects and Sample Weights in Complex Health Survey Data: A Review of Published Articles Using Data From 3 Commonly Used Adolescent Health Surveys. American Journal of Public Health [Internet]. 2012 Jul [cited 2021 Mar 13];102(7):1399–405. Available from: http://ajph.aphapublications.org/doi/10.2105/AJPH.2011.300398

111. Mice: An approach to sensitivity analysis [Internet]. [cited 2020 Jul 12]. Available from: https://www.gerkovink.com/miceVignettes/Sensitivity_analysis/Sensitivity_analysis.html

112. Hanson TE. Inference for mixtures of finite polya tree models. Journal of the American Statistical Association [Internet]. 2006;101(476):1548–65. Available from: https://doi.org/10.1198/016214506000000384

113. Choi S, Cho H. Journal of the Korean Statistical Society Accelerated failure time models for the analysis of competing risks. Journal of the Korean Statistical Society [Internet]. 2019;48(3):315–26. Available from: https://doi.org/10.1016/j.jkss.2018.10.003

114. Jin Z. Semiparametric accelerated failure time model for the analysis of right censored data. Communications for Statistical Applications and Methods [Internet]. 2016 Nov [cited 2021 Aug 6];23(6):467–78. Available from: http://www.csam.or.

kr/journal/view.html?doi=10.5351/CSAM.2016.23.6.467

Appendix A: Plagiarism

Declaration



PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Morelearnings Sibanda (Student number: 1620303) am a student registered for the degree of Msc Biostatistics in the academic year 2021.

I hereby declare the following:

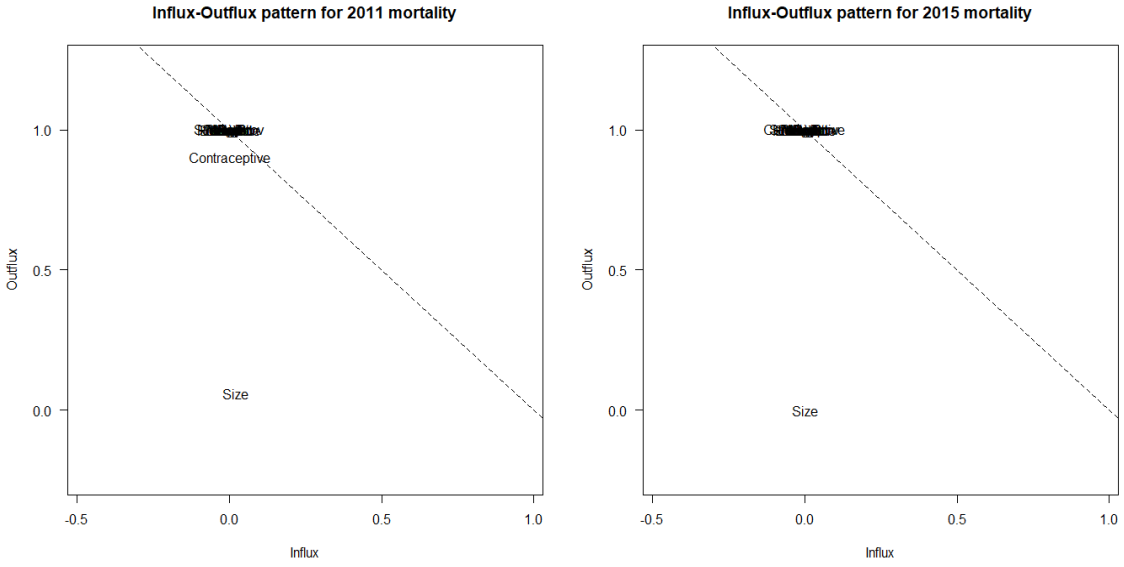
- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: _____ Date: 28-04-2021

Appendix B: Missing Data

Sensitivity Analysis Of Missing Data

Before the imputation of missing values, we analysed how sensitive our data was to imputation and how much influence the imputed models have on the final analysis. To do this we start of by observing the flux plots for data. Flux is used to measure the predictive power of the variables with missing values through the outflux and the dependency on imputed values through the influx (111). Complete variables have an outflux of 1 and an influx of 0.0, that is high predictive power and no dependence on imputation.

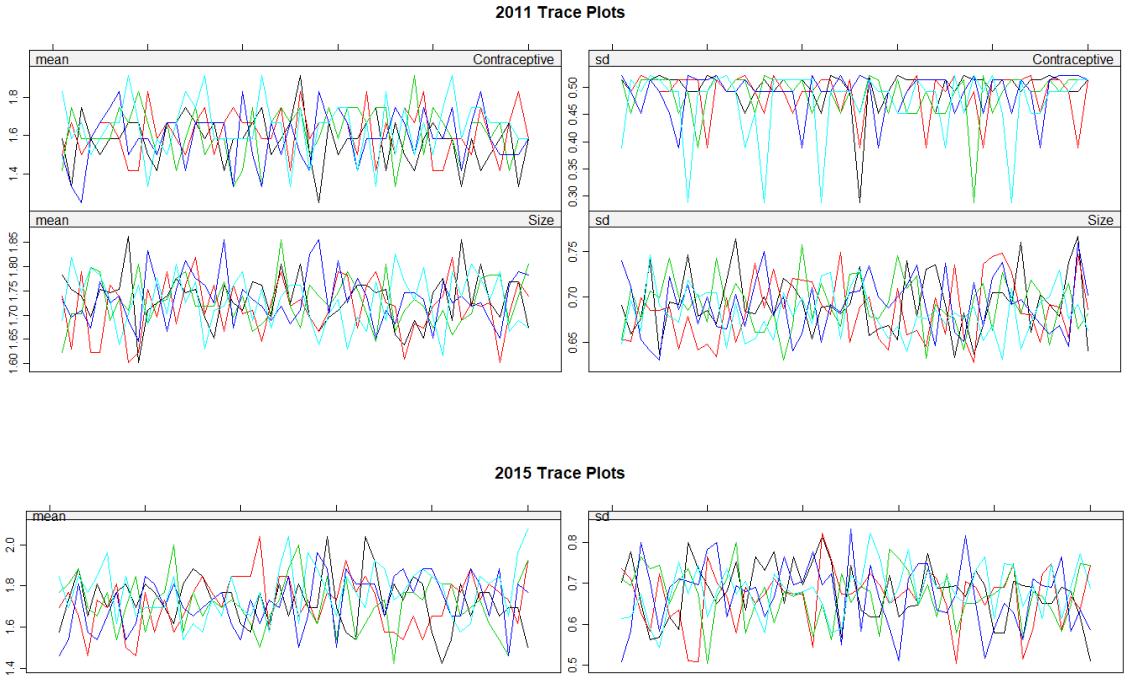


Contraceptive use had a high predictive power in 2011 data and a very low influx and size of baby at birth had low outflux and low influx of 0.023 meaning that both

variables did not depend on the imputation models. In 2015 data, size of baby at birth had a low percentage of missing values but it's the only variable with incomplete data and hence it has the lowest predictive power but non the less, it did not depend on the imputed values shown by a very low influx of 0.004. We decided to take a cautious approach in our imputation for the sake of WinBugs since these variables had negligible dependence on imputed values.

Diagnostics Plots of Missing Data

Trace Plots



Healthy mixing of 5 chains of the mice algorithm showing convergence.

Appendix C: Model Specification in R

Bayesian spatial modeling in R statistical software was perfomed using the *spBayesSurv package* developed by Zhou et al. to fit spatial survival models(47).The implementation of the Cox PH and AFT models in *spBayesSurv* is very similar, the major differences are in the distribution ($f(.)$) and the survival ($S(.)$) functions.

Cox-PH Model

The Cox proportional hazard model incorporated spatial random effects which allow for heterogeneity of the hazard (112) . The *spBayesSurv package* was used in conjunction with the survival and R2BayesX packages for spatial survival analysis in R statistical software. The package requires that a baseline functional form be defined for the Cox-PH model full functionality and the default log-logistic baseline distribution was adopted.

The survival and hazard equations are given below:

$$S(t) = S_0(t)^{e^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i}} \tag{1}$$

$$f(t) = e^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i} S_0(t)^{e^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i-1}} f_0(t) \tag{2}$$

Accelerated Failure Time model (AFT)

For the AFT models, the package allows for three types of AFT models, namely log-logistic, log-normal and Weibull defined according to the baseline survival distribution. The AFT model allows the covariates to be a multiplicative function on time, that is the regression coefficients (β) quantify the impact of explanatory variables (X) on the survival time (113,114). The survival and hazard equations are given below:

$$S(t) = S_0(te^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i}) \tag{3}$$

$$f(t) = e^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i} f_0(e^{\mathbf{X}_{ij}^T\boldsymbol{\beta}+v_i}t) \tag{4}$$

Where:

$i = 1, \ldots, m$ is the number of districts in Zimbabwe,

$j = 1, \ldots, n_i$ are individuals in district i such that $\sum_{i=1}^{m} n_i$ is the total population,

$\mathbf{X}_{ij}$ represents the vector of covariates,

t_{ij} is the survival time for individual j from district i,

β are the regression coefficients

v_i represents the district level frailties,

$S_0(t)$ is the baseline survival,

$f_0(t)$ is the probability density function of the baseline survival.

The Likelihood

Let the data be $D = \{\mathbf{X}_{ij}, t_{ij}\}$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$. The likelihood is given by:

$$L(D \mid \mathbf{W}_L, \theta, \beta, \mathbf{v}) \quad = \quad \pi_{i=1}^{m} \pi_{j=1}^{n_i} [S(a_{ij}) - S(b_{ij})]^{I(a_{ij} < b_{ij})} f(a_{ij})^{I(a_{ij} = b_{ij})} \tag{5}$$

The Priors

The baseline survival $S_0(.)$ follows a transformed Bernstein polynomial (TBP)

$$S_0(.) \mid \alpha, \theta \quad \sim \quad TBP_L(\alpha, S_\theta(.)) \tag{6}$$

such that $\alpha \sim \Gamma(a_0, b_0)$ and $\theta \sim N(\theta_0, \mathbf{V}_0)$

$$\Rightarrow p(\theta) \quad = \quad \exp\left[\frac{-1}{2}\left(\frac{\theta - \theta_0}{V_0}\right)^2\right] \tag{7}$$

such that $TBP_L(\alpha, S_0())$ is given by:

$$S_0(t) \quad = \quad \sum_{j=1}^{L} W_j I(S_\theta(t) \mid j, L - j + 1) \tag{8}$$

where $\mathbf{W}_L = (W_1, \ldots, W_L)^T$ are weights such that $\mathbf{W}_L \sim Dirichlet(\alpha, \ldots, \alpha)$ and as $\alpha \to \infty$, $S_0(.) \to S_\theta(.)$. $I(. \mid a, b)$ reperesents the beta cumulative density function and $S_\theta(t)$ is the distribution of the baseline hazard which can be log-normal, log-logistic or Weibull.

Distribution	$S_\theta(t)$
Log-Logistic	$1 + (e^{\theta_1})^{e^{\theta_2}}$
Log-Normal	$1 - \phi[(log(t) + \theta_1)e^{\theta_2}]$
Weibull	$1 - [-(e^{\theta_1}t)^{e^{\theta_2}}]$

The regression coefficients (β) follow a normal distribution with mean β_0 and variance σ^2; $\beta \sim N(\mu, \sigma_2)$

$$\Rightarrow p(\beta) = \exp\left[\frac{-1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2\right] \tag{9}$$

The unstructured random effect (u) were assumed to follow a Gaussian distribution with mean = 0 and precision τ_u; u $N(0.0,\tau_u)$ (Equation 2.25) with the precision parameter assigned a Gamma hyper-prior (Equation 2.38). The structured spatial effects v were assumed to follow an ICAR prior (Equation 2.35) with the corresponding hyper-prior assumed to follow a Gamma distribution (Equation 2.39).

The Posterior

The Posterior is given by: Posterior $\propto$ Likelihood $\times$ Prior

$$\begin{aligned} P(\mathbf{W}_L, \theta, \beta, v \mid D) &= L(D \mid, \theta, \beta, v)P(\theta)P(\beta)P(v \mid \tau_v)P(\tau_v) \\ &= \Pi_{i=1}^{m}\Pi_{j=1}^{n_i}[S(a_{ij}) - S(b_{ij})]^{I(a_{ij}<b_{ij})} f(a_{ij})^{I(a_{ij}=b_{ij})} \\ &\times \exp\left[\frac{-1}{2}\left(\frac{\theta-\theta_0}{V_0}\right)^2\right] \\ &\times \exp\left[\frac{-1}{2}\left(\frac{\beta-\mu}{\sigma}\right)^2\right] \\ &\times \exp\left[\frac{-\tau_v}{2}\sum_{i\sim j}(v_i - v_j)^2\right] \times \tau_v^{a_v-1}e^{-b_v\tau_v} \end{aligned} \tag{10}$$

In the *spBayesSurv package* in R, there is no way of having both the unstrustured and structured effects in the same model, that is there is no convoolution model, so the spatially correlated and uncorrelated models are run seperately. For spatially uncorrelated model, we replace $P(v \mid \tau_v)P(\tau_v)$ in posterior by

$$P(u \mid \tau_u)P(\tau_u) \quad = \quad \exp\left[\frac{-1}{2}u^2\tau\right]\tau_u^{a_u-1}e^{-b_u\tau_u} \tag{11}$$

The R package uses a Metropolis-Hastings MCMC algorithm for sampling (47). A burn-in of 1000 and thinning after every 5 iterations were set to get a sample of 4,000 from 21,000 iterations. The models were run using less informative and more informative priors, which respectively took values ($\tau \sim \Gamma(0.001, 0.001)$) and ($\tau \sim \Gamma(0.1, 0.1)$).

Appendix D: Model Selection

Cox PH Model Selection

Year	Prior	R DIC	Openbugs DIC
2011	$\Gamma(0.001, 0.001)$	6582	4809
	$\Gamma(0.1, 0.1)$	6577	4818
2015	$\Gamma(0.001, 0.001)$	5887	4555
	$\Gamma(0.1, 0.1)$	5878	4552

The DICs were used for model selection and the models with the lowest DIC values were selected. Models with more informative priors $\Gamma(0.1, 0.1)$, were selected for both survey years in R and for 2015 survey in Openbugs while the model with a less informative prior $\Gamma(0.001, 0.001)$, was selected for 2011 in Openbugs.

AFT Model Selection

Year	Model	Prior	R DIC	Openbugs DIC
2011	Log-Logistic	Γ(0.001,0.001)	6585	6579
		Γ(0.1,0.1)	6584	6571
	Log-Normal	Γ(0.001,0.001)	6570	6562
		Γ(0.1,0.1)	6572	6558
	Weibull	Γ(0.001,0.001)	6588	6594
		Γ(0.1,0.1)	6590	6586
2015	Log-Logistic	Γ(0.001,0.001)	5906	5919
		Γ(0.1,0.1)	5908	5913
	Log-Normal	Γ(0.001,0.001)	5892	5899
		Γ(0.1,0.1)	5894	5898
	Weibull	Γ(0.001,0.001)	5909	5925
		Γ(0.1,0.1)	5908	5922

The log-normal parametric models performed better than the log-logistic and weibull models for both survey years and for the two statistical softwares(R and OpenBugs), hence log-normal models were selected for interpretation. For the 2011 dataset, OpenBugs log-normal model with more informative prior had the lowest DIC whereas for the 2015 dataset, the R log-normal model with less informative prior had the lowest DIC.

Appendix E: Model Codes

OpenBugs Model Codes

Cox-PH Codes

```
#COX PH
model {
for(i in 1:N) {
   for(j in 1:T) {
      Y[i,j] <- step(obs.t[i] - t[j] + eps)
      dN[i, j] <- Y[i, j] * step(t[j + 1] - obs.t[i] - eps) * fail[i]
      }
   }
for(j in 1:T) {
   for(i in 1:N) {
      dN[i, j] ~ dpois(Idt[i, j])
      Idt[i, j] <- Y[i, j] * exp( beta[1]*Twin[i] + beta[2]*Residence[i] +
beta[3]*Sex[i] + beta[4]*Contraceptive[i] + beta[5]*Age[i] +
beta[6]*equals(Religion[i], 1) +  beta[7]*equals(Religion[i], 2) +
beta[8]*equals(Religion[i], 4)  +  beta[9]*equals(Wealth[i], 2)  +
beta[10]*equals(Wealth[i], 3)  +  beta[11]*equals(Married[i], 0) +
beta[12]*equals(Married[i], 2) +  beta[13]*equals(Education[i], 1) +
beta[14]*equals(Education[i], 2) + beta[15]*equals(Size[i], 1) +
beta[16]*equals(Size[i], 3) +u[ID_2[i]] + v[ID_2[i]])*dL0[j]
      }
   dL0[j] ~ dgamma(mu[j], c)
   mu[j] <- dL0.star[j] * c
   }
c <- 0.001
r <- 0.1
for (j in 1 : T) {
   dL0.star[j] <- r * (t[j + 1] - t[j])
   }
# Beta  priors
for(i in 1:16){
beta[i]~dnorm(0.0,0.0001)
}
# Normal  prior distribution for unstructured random effects:
for(k in 1 : Npairs) {
        u[k] ~ dnorm(0.0, tauu);
        }
# CAR prior distribution for structured random effects:
          v[1 : Npairs] ~ car.normal(adj[], weights[], num[], tauv)

for(k in 1:sumNumNeigh) {
             weights[k] <- 1
                    }
# Hyper-Priors :
          tauv  ~ dgamma(0.1, 0.1)            # prior on precision
          tauu ~ dgamma(0.1, 0.1)
```

AFT Codes

```
# AFT MODEL
model {
for(i in 1:N) {
#temp denotes(logt - beta0 - beta*x - frailty)/sigma1
temp[i] <- (log(obs.t[i] + 0.01)  - beta0  - beta[1]*Twin[i] -
beta[2]*Residence[i] - beta[3]*Sex[i] - beta[4]*Contraceptive[i]   -
beta[5]*Age[i] - beta[6]*equals(Religion[i], 1) -
beta[7]*equals(Religion[i], 2) -  beta[8]*equals(Religion[i], 4)  -
beta[9]*equals(Wealth[i], 2)  - beta[10]*equals(Wealth[i], 3)  -
beta[11]*equals(Married[i], 0) - beta[12]*equals(Married[i], 2) -
beta[13]*equals(Education[i], 1) - beta[14]*equals(Education[i], 2) -
beta[15]*equals(Size[i], 1) - beta[16]*equals(Size[i], 3) - u[ID_2[i]] -
v[ID_2[i]])/sigma

#SURVIVAL DISTRIBUTION AND DENSITY FUNCTION OF
STANDARD NORMAL
 s0[i] <- 1 - phi(temp[i])
 f0[i] <-  pow(2*3.14, - 0.5) * exp(-0.5*pow(temp[i], 2))
# LOG - LIKELIHOOD FUNCTION
L[i] <- fail[i]* log(f0[i]/(sigma*(obs.t[i] + 0.01))) + (1 - fail[i])*log(s0[i])
#POISSON ZERO TRICK
zeros[i] <- 0
new[i] <- -L[i]
zeros[i] ~ dpois(new[i])
}
#CAR MODEL SPECIFICATION
for(j in 1:sumNumNeigh) {weights[j] <- 1}
v[1:Npairs] ~ car.normal(adj[], weights[], num[], tauv)
v.mean <- mean(v[])

#IID NORMAL MODEL SPECIFICATION
for(k in 1:Npairs){
u[k] ~ dnorm(0.0, tauu)
}
#PARAMETER PRIOR FOR THE PARAMETER IN THE AFT
MODEL
beta0 ~ dnorm(0.0, 0.001)
for(i in 1:16){
beta[i] ~ dnorm(0.0, 0.001)
}

inversesigma ~ dgamma(0.001, 0.001)
 sigma <- 1/inversesigma
#PARAMETER IN THE CAR MODEL OR NORMAL MODEL
tauv ~ dgamma(0.1, 0.1)
tauu ~ dgamma(0.1, 0.1)
```

R Model Cox-PH and AFT Codes

```
rm(list = ls())

library("coda")
library("survival")
library("spBayesSurv")
library("fields")
library("BayesX")
library("R2BayesX")
library('unikn')
library("haven")
library("mcmcplots")
library(prettymapr)

##############################COX-PH MODEL#######################################
#for non-spatial replace frailtyprior("car", ID_2) with frailtyprior("iid", ID_2)
set.seed(123)
mcmc <- list(nburn = 2000, nsave = 2500, nskip = 4, ndisplay = 100)
prior <- list(maxL = 15, taua0 = 0.001,taub0 = 0.001)
ptm <- proc.time()

res15_1 <- survregbayes(formula = Surv(Survtm,Dead) ~ Age_Cat + Married + Residence +
                          Education + Religion + Wealth
                        + Contraceptive + Twin + Sex + Size
                        + frailtyprior("car", ID_2), data = d, survmodel = "PH",
                        dist = "loglogistic", mcmc = mcmc, prior = prior, Proximity = E )


##############################AFT MODEL#######################################
#for non-spatial replace frailtyprior("car", ID_2) with frailtyprior("iid", ID_2)

set.seed(456)
mcmc <- list(nburn = 2000, nsave = 2500, nskip = 4, ndisplay = 100)
prior <- list(maxL = 15, taua0 = 0.1,taub0 = 0.1)
ptm <- proc.time()

res1_1 <- survregbayes(formula = Surv(Survtm,Dead) ~ Age_Cat + Married + Residence +
                         Education + Religion + Wealth
                       + Contraceptive + Twin + Sex + Size
                       + frailtyprior("car", ID_2), data = d, survmodel = "AFT",
                       dist = "loglogistic", mcmc = mcmc, prior = prior, Proximity = E )
```

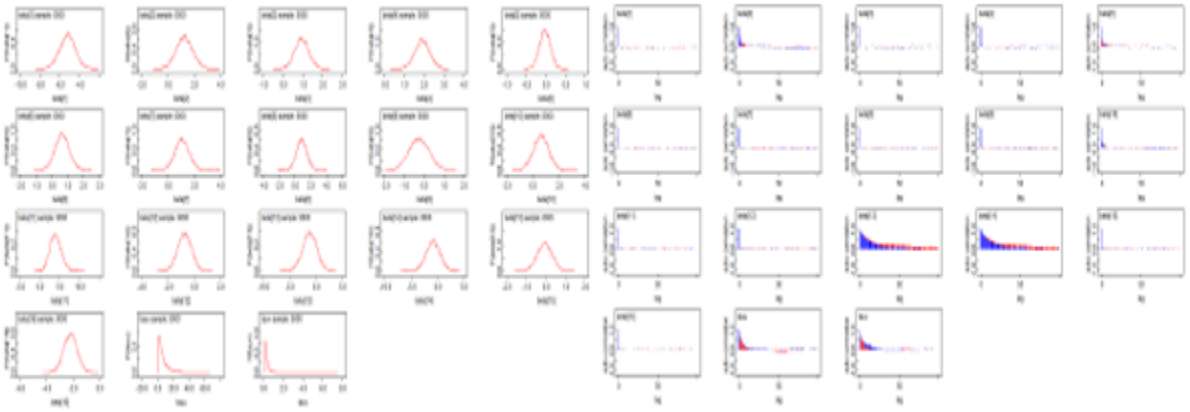
Appendix F: Diagnostics Plots

OpenBugs Diagnostics Plots

2011 Diagnostics Plots

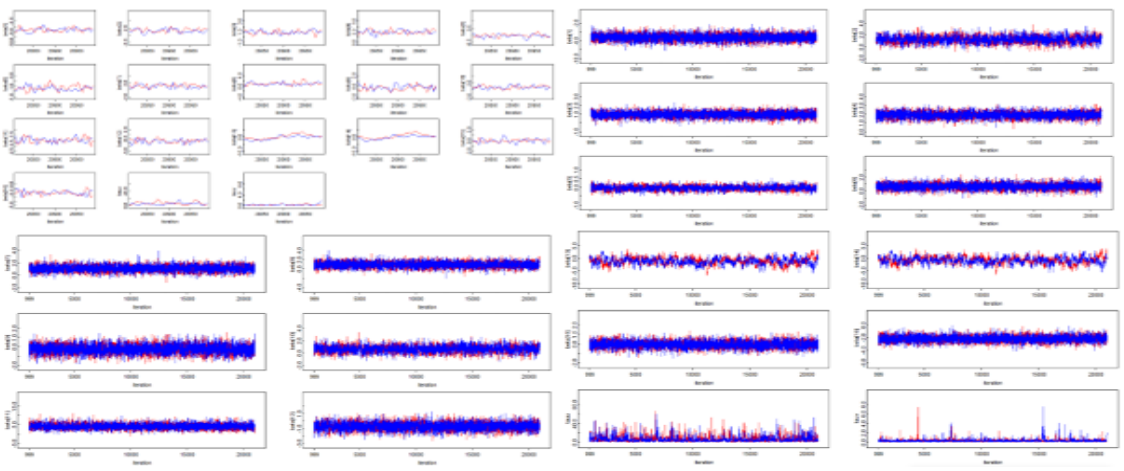
Density Plots

Auto-Correlation Plots



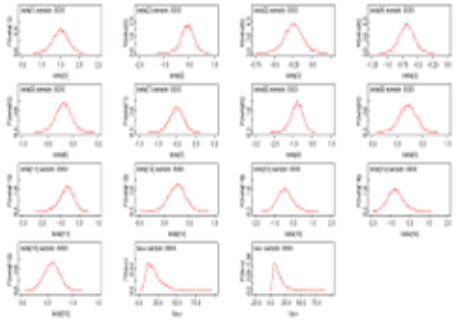
Trace Plots

Time Series Plots

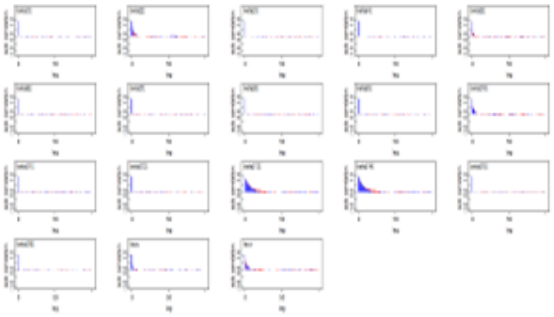


2015 Diagnostics Plots

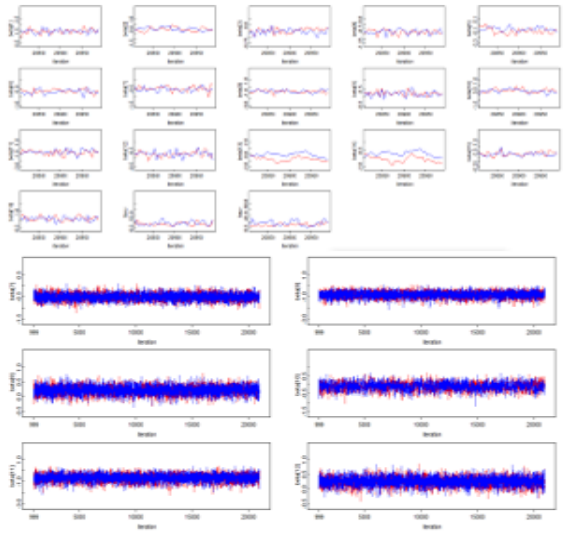
Density Plots



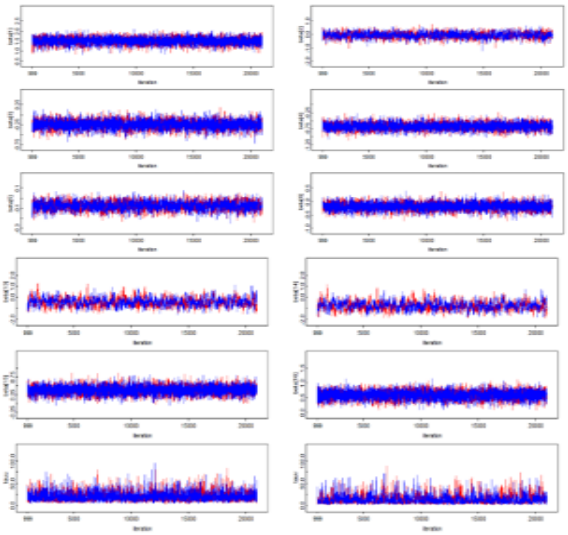
Auto-Correlation Plots



Trace Plots



Time Series Plots



R Diagnostics Plots

2011 Diagnostics Plots



2015 Diagnostics Plots



Appendix G: Ethics Certificate

.

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

04/09/2020

Ref: W-CBP-200904-01

TO WHOM IT MAY CONCERN:

Waiver: This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical).

Investigator: Mr M Sibanda
Student No. (if appropriate): 1620303
Staff No. (if appropriate):

Supervisor: Professor NB Kandala

School: Public Health
Department: Epidemiology and Biostatistics
Medical School
University

Project title: *Spatial modelling of under-5 child mortality and associated factors in Zimbabwe using 2011 and 2015 demographic and health survey data*

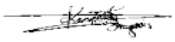
Reason: Review of information in the public domain.
No human participants will be involved in the study.

Dr CB Penny
Co-Chairperson: Human Research Ethics Committee (Medical)

Research Office Secretariat:
Physical address: Phillip Tobias Building, 3rd Floor, Office 302, Corner York Road and Princess of Wales Terrace, Parktown, Johannesburg 2193.
Postal address: Private Bag 3, Wits 2050
Tel Nos. +27 (0)11-717-1234/2656/2700/1252
Office E-mail: HREC-Medical.ResearchOffice@wits.ac.za.
Website:http://www.wits.ac.za/research/about-our-research/ethics-and-research-integrity/

Appendix H: Turnitin Report

.



1620303:1620303Plag.pdf

ORIGINALITY REPORT

13%	11%	10%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	hdl.handle.net Internet Source	2%
2	www.wits.ac.za Internet Source	1%
3	journals.plos.org Internet Source	1%
4	bmcpublichealth.biomedcentral.com Internet Source	1%
5	www.researchsquare.com Internet Source	<1%
6	Ngianga-Bakwin Kandala, Gebrenegus Ghilagaber. "A Geo-Additive Bayesian Discrete-Time Survival Model and its Application to Spatial Analysis of Childhood Mortality in Malawi", Quality & Quantity, 2006 Publication	<1%
7	David G. Kleinbaum, Mitchel Klein. "Survival Analysis", Springer Science and Business Media LLC, 2012 Publication	<1%