Trends in the Under-five Mortality and Disparities between Regions in Kenya

Shikuku, Eric Odongo

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Statistical Science of Strathmore University

Strathmore Institute of Mathematical Sciences

Strathmore University

Nairobi, Kenya

April 2025

Declaration

I declare that this work has not been previously submitted and approved for award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Name: **Shikuku Eric Odongo**

Signature:

Date: April 1, 2025

Approval

The thesis of Shikuku Eric Odongo was reviewed and approved by the following:

Dr. Thaddeus Egondi

Supervisor, Thaddeus Egondi

Institute of Mathematical Sciences, Strathmore University. 01 April 2025

Prof. Bernard Omolo

Supervisor, Oguna-Omolo

Institute of Mathematical Sciences, Strathmore University.

Dr. Godfrey Madigu

Dean,

Institute of Mathematical Sciences, Strathmore University.

Dr. Bernard Shibwabo

Director,

Office of Graduate Studies, Strathmore University.

Abstract

This study examined trends and regional disparities in under-five mortality in Kenya using reconstructed data from the Kenya Demographic and Health Surveys (KDHS) spanning 2000 to 2022. The primary objectives were to quantify the differences in child survival across Kenya's regions and to identify key socio-economic, demographic, and healthcare-related determinants influencing these disparities.

Survival analysis using Kaplan–Meier estimation and Cox proportional hazards models revealed that under-five mortality rates have declined significantly over time – from approximately 33 deaths per 1,000 live births in the 1999–2003 period to 11.6 per 1,000 in 2018–2022. However, substantial regional differences persist: regions such as Nyanza and Western exhibited the poorest survival outcomes with mortality rates as high as 27.6 per 1,000, whereas Central and Rift Valley regions showed the lowest mortality rates (approximately 13 per 1,000). Stratification by maternal education further indicated that children born to mothers with higher education levels experienced better survival probabilities compared to those with lower education, and survival analyses by sex demonstrated that female children had a consistently lower risk of mortality (adjusted hazard ratio [HR] = 0.81, $p < 0.001$) relative to males.

The adjusted Cox regression model, which controlled for wealth index, maternal age, smoking status, type of residence, and preceding birth interval, highlighted that being in the richest wealth category was paradoxically associated with a 41% higher hazard of mortality compared to the poorest group ($p = 0.004$), suggesting potential inequities in healthcare access or reporting. Other determinants, including maternal age and smoking status, were not found to be statistically significant in the adjusted analysis, while preceding birth interval exhibited a borderline effect.

KEY WORDS: mortality of under five years of age, censored, event, Cox proportional hazards model.

Table of contents

Acknowledgement

First and above all, I am grateful to God for His provision, and grace, and for granting me good health throughout the study period.

I am grateful to my family for their moral support.

I am indebted to my supervisors, Dr. Thaddaeus Egondi and Prof. Bernard Omolo for their availability, kind guidance, and support without which this proposal would not have been a success.

Chapter 1

Introduction

Under-five mortality is a significant public health concern in most countries, including Kenya, and therefore, proper intervention policies need to be implemented to reduce infant mortality rates. This can only be done once disparities between regions and trends have been identified. The main focus of this research was to determine these trends and regional disparities.

According to what has been published, various factors are associated with variations in mortality among under-fives. These consist of individual characteristics such as nutrition, wealth index, and levels of education (Misselhorn and Harttgen, 2006). Other factors associated with child mortality include, antenatal care, breastfeeding duration, maternal health care services, birth spacing, and the mother's age (Adetunji, 1995). These, among other newly identified factors (eg vaccination status, diarrhea, smoking status, and distance to the nearest health facility), formed part of the independent variables in the regression models that were considered in this research to account for disparities.

According to the fourth-millennium development goal, it is expected that all member countries of the United Nations should have attained a reduction in infant mortality rates by two-thirds by 2015 (Journals, 2020).

Kenya is among the countries still fighting to reduce infant mortality rates and achieve this Millennium Development Goal, which is overdue and now aims at the 2030 target of 25 deaths per 1000 live births (Keats et al., 2018).

1.1 Background to the Study

Time-to-event analysis or survival analysis is widely used when the outcome/dependent variable relates to the time taken until the occurrence of a pre-specified event. These methods consider censoring mechanisms to ensure accurate analysis without discarding information on participants who fail to experience the event of interest within the study period. The event of interest could be time until death, time to the first onset of an adverse event, time to relapse among drug addicts, etc. This study focused on analyzing the time taken until the death of an infant due to the top five major causes, which include preterm birth complications, low birth weight, intrapartum growth restrictions, neonatal infections, and congenital malformations (Imbo et al., 2021).

1.2 Statement of the Problem

Despite various public health interventions aimed at reducing the mortality of children below five years, under-five mortality remains a significant public health challenge in Kenya. Although numerous studies have analyzed the determinants of mortality for children under five years of age, there is a gap in understanding regional disparities in these mortality rates between the eight regions of the country and regional trends since 2000. The regions were based on former administrative units (provinces), these are Nairobi, Central, Nyanza, Eastern, North Eastern, Coastal, and Rift Valley.

This research sought to fill this gap by looking at the disparities between regions and trends by reconstructing Demographic Health Surveys from 2000 to 2022. These findings aimed to inform policy and improve public health interventions to reduce regional disparities and achieve major reductions in mortality rates of under-five years of age.

1.3 Research Objectives

The study's main objective was to examine the disparity between regions and the yearly trends in the mortality rates of the under-fives since 2000.

The specific objectives of this research were;

i) To examine the extent of regional disparities in the trends in child mortality in Kenya using data from the DHS survey.

ii) To identify key socio-economic, demographic, and healthcare-related determinants contributing to regional differences in child mortality trends.

1.4 Research Questions

i) What is the extent of regional disparities in trends in child mortality rates?

ii) What are the key social-economic, demographic, and healthcare-related determinants that contribute to regional differences in child mortality trends?

1.5 Justification of the Study

This study was helpful because it highlighted how various regions of the country are performing in terms of reducing the national burden of under-five mortality. Variations in known factors associated with mortality among children aged five years and under in regions will help to develop appropriate public health intervention mechanisms. The yearly trends were useful in knowing whether these rates are declining or increasing over time and whether these trends are significant or not.

1.6 Significance of the Study

The study was significant in multiple dimensions, particularly in public health, policy formulation, and statistical methodology:

Public Health Impact

Identifying regional disparities in under-five mortality rates provided critical insights into how different regions of Kenya are progressing in reducing child mortality. Having an understanding of the critical factors associated with disparities in under-five mortality helped target effective interventions and optimized resource allocation to areas with the highest burden. The findings also informed strategies to improve maternal and child health services, aligning efforts to meet Sustainable Development Goals (SDG), especially SDG 3 (Good Health and Well-being).

Policy and Planning

The study provided evidence for policymakers to design tailored interventions addressing regional inequalities in child survival outcomes. By pinpointing significant factors associated with mortality among under-fives, research helped prioritize actionable factors, such as access to healthcare care, socioeconomic determinants, and maternal education.

Statistical and Methodological Significance

Survival analysis techniques, such as Cox proportional hazards, improve the understanding of time-to-event data in child mortality research. The evaluation of the performance of the model (residuals) and the assumptions (proportional hazards) strengthened the robustness of the conclusions, setting a methodological benchmark for future research.

Chapter 2

Literature Review

2.1 Introduction

The study of under-five mortality is crucial in understanding health disparities and identifying effective interventions. Various studies have explored child mortality trends, determinants, and regional differences. However, gaps remain in understanding the temporal and spatial disparities in under-five mortality rates across Kenyan regions. This chapter reviewed existing literature on under-five mortality, focusing on statistical methodologies such as survival analysis, Cox regression models, trend analysis, and the calculation of mortality rates. The review also discussed Kenya's current situation regarding under-five mortality, highlighting key interventions and their effectiveness.

2.2 A Review of Existing Literature

2.2.1 Trends in Under-Five Mortality

Under-five mortality has shown significant variations across different regions and periods. Kimani-Murage et al. (2014) analyzed trends in child mortality in Kenya and found a general decline over three decades. However, disparities between rural and urban areas persisted, with rural regions experiencing slower mortality reductions. Similar studies in Uganda and East Africa (Nasejje et al., 2015) and (Tesema et al., 2021) confirmed these patterns, indicating that socioeconomic and healthcare access differences drive regional disparities.

Macharia et al. (2019) examined under-five mortality trends in Kenya since 1965 and found a 61.6% national decline. However, county-level variations were significant, with some regions showing mortality rates nearly four times higher than others. This underscores the need to explore disparities at the regional level. The current study built upon those findings by reconstructing yearly trends from 2000 to 2022 to understand regional disparities better.

Studies have also analyzed the impact of global health initiatives, such as the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs), on child mortality rates. Keats et al. (2018) highlighted how increased funding for maternal and child health services, improved vaccination coverage, and enhanced sanitation have contributed to mortality declines. However, structural inequalities and access disparities continue to hinder uniform progress.

2.2.2 Determinants of Under-Five Mortality

Several factors contribute to under-five mortality, including socioeconomic status, healthcare accessibility, maternal education, and environmental conditions. Victora et al. (2003) emphasized the role of poverty in child mortality, noting that children from low-income families experience higher mortality rates due to inadequate nutrition and healthcare access. Similar findings were reported by (Ettarh and Kimani, 2012), who found that rural Kenyan children face higher mortality risks than their urban counterparts.

Bell et al. (2010) explored social determinants of health, indicating that factors such as maternal education, antenatal care, and vaccination programs significantly impact child survival. Tesema et al. (2021) used a multilevel Weibull regression model to identify key determinants of under-five mortality in East Africa. Their findings highlighted the importance of maternal education, wealth index, and antenatal care in reducing mortality rates.

Environmental factors, such as access to clean drinking water and proper sanitation, also play a significant role in under-five mortality. Macharia et al. (2021) found that regions with inadequate sanitation facilities and high pollution levels reported significantly higher mortality rates. This aligns with findings from the World Health Organization (WHO), which stresses the importance of environmental health interventions in reducing child mortality.

This study extended what the above authors have published by identifying regional disparities in these determinants, providing a more localized perspective on factors influencing child survival.

2.2.3 Statistical Methods for Analyzing Under-Five Mortality

Cox Proportional Hazards Model

Survival analysis techniques are commonly used to study under-five mortality due to the presence of censored data (children who survive beyond five years). The Cox proportional hazards model is particularly useful as it estimates hazard ratios without assuming a specific distribution for survival times (Cox, 1972). Nasejje et al. (2015) applied this model in Uganda, revealing significant household-level factors influencing child mortality.

A major assumption of the Cox model is the proportional hazards assumption, which can be tested using Schoenfeld residuals. If this assumption holds, the model provides reliable estimates of risk factors affecting child survival (Cleves, 2008). This study employed the Cox regression model to analyze survival probabilities and identified determinants of under-five mortality across Kenyan regions.

Kaplan-Meier Estimation and Log-Rank Tests

The Kaplan-Meier estimator is widely used for non-parametric survival analysis. It provides survival probability estimates over time and allows comparisons across different groups (Kaplan and Meier, 1958). The log-rank test statistically compares survival distributions between groups, making it useful for assessing regional disparities in under-five mortality (Macharia et al., 2021).

Trend Analysis and Time Series Methods

Trend analysis is essential for understanding mortality patterns over time. Common methods include:

- Joinpoint Regression: Identifies significant changes in trends by fitting multiple linear segments to time-series data (Kim et al., 2000).
- ARIMA Models: Used for forecasting mortality rates by analyzing temporal dependencies in time-series data (Box et al., 2015).
- Bayesian Time-Series Models: Incorporate prior distributions to estimate mortality trends more accurately, especially in regions with sparse data (Gelman et al., 1995).

These methods allow researchers to detect patterns and predict future trends in child mortality.

2.2.4 Calculation of Mortality Rates

Mortality rates are a fundamental demographic measure used in epidemiology, public health, and actuarial sciences. The calculation of mortality rates allows researchers and policymakers to assess population health, compare trends over time, and develop health interventions. This review summarizes key methodologies and literature on the calculation of mortality rates, including crude mortality rates, age-specific rates, and standardized mortality ratios.

Crude Mortality Rate (CMR)

The crude mortality rate (CMR) is a basic way to assess mortality within a population. It represents the total number of deaths occurring in a specific group over a set period, typically reported per 1,000 or 100,000 people. As noted by (Anderson and Rosenberg, 1998) CMR offers a general overview of public health but has limitations since it does not adjust for variations in age distribution.

Age-Specific Mortality Rate (ASMR)

Age-specific mortality rate (ASMR) accounts for variations in mortality risk across different age groups. The method involves calculating mortality rates within predefined age cohorts. Preston (2000) highlight the importance of ASMR in understanding disease patterns and assessing the impact of aging on population health. Furthermore, ASMR enables researchers to identify high-risk age groups, facilitating targeted health interventions.

Standardized Mortality Ratio (SMR)

The Standardized Mortality Ratio (SMR) is a key measure used to compare mortality rates between populations with different age structures. The SMR is calculated as the ratio of observed deaths in a study population to the expected number of deaths based on a reference population. Breslow et al. (1980) emphasize that SMR helps in adjusting for confounding variables, making it useful for epidemiological studies comparing occupational, regional, or disease-specific mortality rates.

Life Table Analysis

Life table analysis is another widely used method for mortality estimation. It provides a comprehensive view of mortality patterns by estimating survival probabilities at different ages. Chiang (1984) presents a detailed methodology for life table construction, including both cohort and period life tables. Life table analysis is instrumental in actuarial studies, social security planning, and health policy formulation.

Cause-Specific and Infant Mortality Rates

Cause-specific mortality rates focus on deaths attributable to specific diseases or conditions. Doll and Peto (1981) emphasizes that these rates help in identifying epidemiological trends and evaluating disease control measures. Additionally, the infant mortality rate (IMR) serves as a critical indicator of health care quality and socio-economic conditions. UNICEF (2020) highlights IMR as a benchmark for assessing child health and development interventions

Mortality rates are typically expressed as the number of deaths per 1,000 live births. The standard formula is:

$$MR = \frac{D}{B} \times 1000, \tag{2.1}$$

where;

- D is the number of deaths among children under the age of five years of age.
- B is the total number of live births.

Macharia et al. (2019) used this approach to estimate county-level under-five mortality in Kenya. Adjusted mortality rates can be computed using direct and indirect standardization methods, accounting for differences in population structures across regions.

The current study employed these methods to calculate mean region-specific mortality rates and mean mortality rates across analysis periods in order to analyze disparities over time.

Challenges in Mortality Rate Calculation

Despite advancements in mortality estimation methods, several challenges persist. These include data quality issues, underreporting of deaths, and misclassification of causes of death. Hill et al. (2017) discusses the impact of incomplete data and statistical adjustments required for improving mortality estimates in low-resource settings.

2.3 Current Situation in Kenya

The Kenya Demographic and Health Survey (KDHS) 2022 indicates a continued decline in under-five mortality. In 1993, the mortality rate was 96 deaths per 1,000 live births, decreasing to 41 deaths per 1,000 live births in 2022 (Kedogo, 2023). This decline is attributed to improved maternal healthcare, vaccination programs, and better access to medical facilities.

However, disparities remain. Northern Kenya consistently experiences higher mortality rates due to limited healthcare access and poor socioeconomic conditions (Macharia et al., 2021). The 2030 target under the Sustainable Development Goals (SDGs) aims to reduce under-five mortality to 25 deaths per 1,000 live births (Journals, 2020). Achieving this requires targeted interventions addressing regional inequalities.

2.4 Conclusion

The above literature highlighted significant progress in reducing under-five mortality in Kenya and underscored persistent regional disparities. Studies have identified key determinants, including socioeconomic status, maternal education, and healthcare access. Advanced statistical methods, such as Cox regression and time-series analysis, have been instrumental in mortality research.

The current study relied on some of these work by reconstructing mortality trends from 2000 to 2022, focusing on regional disparities and employing advanced survival analysis techniques. The findings will provided valuable insights for policymakers to design targeted interventions aimed at achieving equitable child survival outcomes.

Chapter 3

Methodology

3.1 Research Methodology

3.1.1 Introduction

This research was based on cross-sectional data collected under the Kenya Demographics Health Survey (KDHS) program between 2003 and 2022. The sample size included all children born in the last five years before the data collection dates. The analysis focused on children born within one to five years before each survey date, who were the result of singleton births, and who either survived infancy or did not.

The outcome variable measured the duration from birth to the death of a child under the age of five. If a child passed away within the first five years of life, it was classified as an **event**. Meanwhile, children who remained alive throughout the study period were treated as **censored** observations.

3.1.2 Research Design

Research Approach

This study employed a quantitative research approach using secondary data for analysis. The research utilized data from the Kenya Demographic and Health Surveys (KDHS) from 2003 to 2022 to analyze trends and disparities in under-five mortality across different regions in Kenya.

Research Design

A longitudinal study design was used to examine historical trends and regional disparities in under-five mortality over time. This design was appropriate as it allowed for an in-depth analysis of time-based trends and their associated determinants.

Study Population and Sample

The study population comprises children under five years whose information is captured in the KDHS datasets from 2000 to 2022. Any child born five years prior to the interview date/year. The sample included all singleton births occurring within one to five years before each survey period. The dataset included censored observations (children who survived beyond five years) and event observations (children who died before reaching five years of age).

Data Collection and Sources

Secondary data was sourced from KDHS surveys conducted between 2003 and 2022. The datasets contained demographic, socioeconomic, maternal health, and child health indicators. The authorization to access the data was obtained from the DHS program.

Dependent variable

The outcome variable for this research was time until the death of an infant before celebrating their fifth birthday.

Independent variables

The independent variables considered in the Cox regression models are sex of the child, mother's level of education, analysis period, region, smoking status of the mother, wealth index, age of the mother, type of residence, and preceding birth interval in months.

Inclusion and Exclusion Criteria

Inclusion Criteria:

- Children under the age of five whose data was recorded in the KDHS surveys from 2003 to 2022.
- Singleton births within five years before each survey year.
- Mothers who provided complete and reliable information on child mortality, demographic, socioeconomic, and healthcare indicators.
- Data entries with clear and verified records on key independent variables such as maternal education, birth interval, and antenatal care use.

Exclusion Criteria:

- Children older than five years or those whose birth records fall outside the survey's timeframe.
- Multiple births (e.g., twins or triplets) to maintain homogeneity in the survival analysis.
- Cases with incomplete or missing data on key independent and dependent variables. The variable has missing values across all of its observations.
- Records with inconsistencies or errors identified during data cleaning processes.

The inclusion criteria ensured that the sample is representative of the target population while focusing on children at risk of under-five mortality. The exclusion of multiple births reduced variability that might arise from unique birth conditions affecting mortality rates. Furthermore, eliminating incomplete data enhanced the robustness and reliability of the statistical analysis.

Scientific Justification of Sampling Method

The KDHS employs a two-stage stratified sampling design, which ensures representativeness and generalizability of findings to the entire Kenyan population. The first stage involves the selection of clusters (enumeration areas) based on a probability proportional to size sampling technique. The second stage involves a systematic sampling of households within selected clusters (Kenya National Bureau of Statistics, 2022). This method ensures an adequate sample size for reliable estimates of mortality rates in children under five years of age while minimizing selection bias (Croft et al., 2018).

The KDHS data set was appropriate for this study as it provided regularly collected, nationally representative, high-quality data on child health and mortality. Its use allowed for a robust examination of trends over time while maintaining consistency in measurement and methodology across survey waves (DHS Program, 2021). Furthermore, the large sample size ensured statistical power in detecting disparities in mortality rates between regions and over time (Rutstein et al., 2006).

Conceptual Framework and Research Gap

The conceptual framework for this study was based on a multi-variable determinant model that incorporated demographic, socioeconomic, healthcare, and environmental factors that influence mortality among under-fives. The framework hypothesized that socioeconomic disparities, accessibility to healthcare care, and maternal characteristics collectively influence the survival of children under five years of age.

The research gap addressed in this study is the limited focus on regional disparities and the temporal trends in mortality among under-fives in Kenya. Although previous studies have analyzed national mortality trends, fewer studies have explored how these determinants vary between regions over time.

The statistical model

Since our dependent variable was the time until an infant's death, we adopted survival analysis methods. The main survival method used is the **Cox regression model**. This model was used because of its popularity in the literature, and it is also appropriate when supplementary information, e.g. demographic variables, is available. Another reason for choosing the Cox model is that it does not make any assumptions about the distribution of survival times, which lends itself to many real-life applications. All the independent variables are assumed to be measured once at the origin of the study and therefore the current research did not cover methods for handling covariates that vary over time i.e time-dependent covariates. The Cox model does not impose a specific probability distribution model on the survival times, the only assumption it imposes is that of proportional hazard. The proportional hazard assumption was evaluated using the Schoenfeld residual test.

The Cox regression model is semi-parametric because the survival times do not assume any specific probability model. The general form of this model is;

$$h_i(t) = exp(\beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_p x_{pi})h_0(t), \tag{3.1}$$

where $h_0(t)$ is the baseline hazard function at time t while $h_i(t)$ is the hazard function for the i^{th} individual at time t. $\beta = (\beta_1, \beta_2, ..., \beta_P)'$ is the vector of coefficients of the p independent variables in the model.

Non-parametric survival comparisons across variable categories were also conducted using the Kaplan-Meier curves and the log-rank statistical tests. The analysis was performed using R software as the main tool, however, other tools like Python, SAS, Microsoft Excel, and Stata were also used for a deeper exploration of the data.

3.1.3 Specific Statistical methods and their justification

The study is designed to examine the extent of regional disparities in child mortality trends in Kenya using data from the DHS survey. To visually compare the survival experiences between different regions, stratified Kaplan-Meier curves will be employed. This method is particularly useful because Kaplan-Meier curves provide a clear visual representation of survival experiences across regions. In addition, stratified log-rank tests will be used to statistically compare these survival experiences by employing a p-value approach, thereby offering a quantitative measure of the differences observed between regions.

Furthermore, the study aims to identify the key socio-economic, demographic, and healthcare-related determinants that contribute to the regional differences in child mortality trends. To achieve this, a Cox Proportional Hazards regression model will be applied.

This model is chosen for its effectiveness in estimating the hazard ratio for each determinant while accounting for censoring in the data. Cox regression is widely recognized for its suitability in handling censored survival data that satisfy the proportional hazards assumption, and it yields interpretable hazard ratios. This method is particularly practical when supplementary information, such as demographic variables, is available, and it does not require the assumption that survival times follow a specific probability distribution.

3.2 Dissemination and Utilization of study results

The results will be distributed through the Strathmore University website and the Strathmore University library, authorized by the administrator of SU @Strathmore University.

This proposal takes into account the welfare, rights, perceptions, and cultural heritage of the participants. The results will be used consistently to ensure maximum benefit for the government, individuals, and other interested parties such as organizations that support public health initiatives.

3.3 Ethical consideration

The approval of the Ethics and Review Committee was obtained before this research was conducted (See Appendix B). This helped ensure compliance with data protection laws for study subjects and international laws on data usage and research.

Chapter 4

Results and Interpretation

4.1 Introduction

This chapter presents the descriptive statistics, diagnostic tests, and results from the fitted models, along with their interpretation. The analysis is based on pooled data from four survey panels conducted in 2003, 2008–2009, 2014, and 2022. The primary outcome variable is the age at death (in months), with the study focusing exclusively on single births to ensure consistency in the analysis.

The primary statistical approach employed is the Cox proportional hazards regression model, which assesses the relationship between explanatory variables and the risk of death over time. A key assumption of this model is the proportional hazards assumption, which will be tested to ensure its validity in this context.

To provide a comprehensive understanding of the dataset, the chapter begins with an exploration of baseline characteristics, including descriptive statistics and visualizations of survival curves. Key categorical variables stratify these survival curves to highlight potential disparities in child mortality risks across different subgroups. Subsequent sections will delve into the results of the statistical models, discussing the significance and implications of key covariates while assessing the robustness of the findings through diagnostic tests.

4.2 Descriptive Statistics and Visualizations

Table 4.1: Descriptive Statistics by Gender

Characteristic	**Male N = 60,250**[1]	**Female N = 58,562**[2]
Time to death	23 (13, 29)	23 (13, 29)
Censoring Status	1,051 (1.7%)	867 (1.6%)
Wealth Index		
Poorest	19,797 (33%)	19,415 (33%)
Poorer	9,710 (16.1%)	9,506 (16.2%)
Middle	10,064 (16.7%)	9,814 (16.8%)
Richer	11,806 (19.6%)	11,039 (18.9%)
Richest	8,788 (14.6%)	8,788 (14.6%)
Marital Status		
Never married	3,742 (6.2%)	6,730 (11.5%)
Married	47,892 (79.5%)	33,733 (57.6%)
Living together	7,800 (12.9%)	9,079 (15.5%)
Widowed	554 (0.9%)	1,744 (3.0%)
Divorced	220 (0.4%)	436 (0.7%)
Not living together	31 (0.1%)	88 (0.2%)
Missing	23,911	23,611
Smoking Status		
No	2,050 (11%)	20,150 (34.4%)
Yes	16,982 (89%)	38,332 (65.6%)
Missing	41,218	80
Type of Residence		
Urban	18,689 (31%)	19,148 (33%)
Rural	41,561 (69%)	39,414 (67%)
Vaccination Status		
No	2,294 (11%)	1,326 (2.3%)
Yes	17,389 (83%)	16,669 (29%)
Missing	40,567	40,567

[1] Mean (SD, Q2), n(%)
[2] Mean (SD, Q2), n(%)

Table 4.1 above shows the frequency counts for the main categorical variables in the data set and the mean, standard deviation, and 50th percentiles for continuous variables such as duration of breastfeeding. It gives us insight into how these variables are distributed among male and female children in our sample.

4.3 Survival Experience stratified by analysis periods

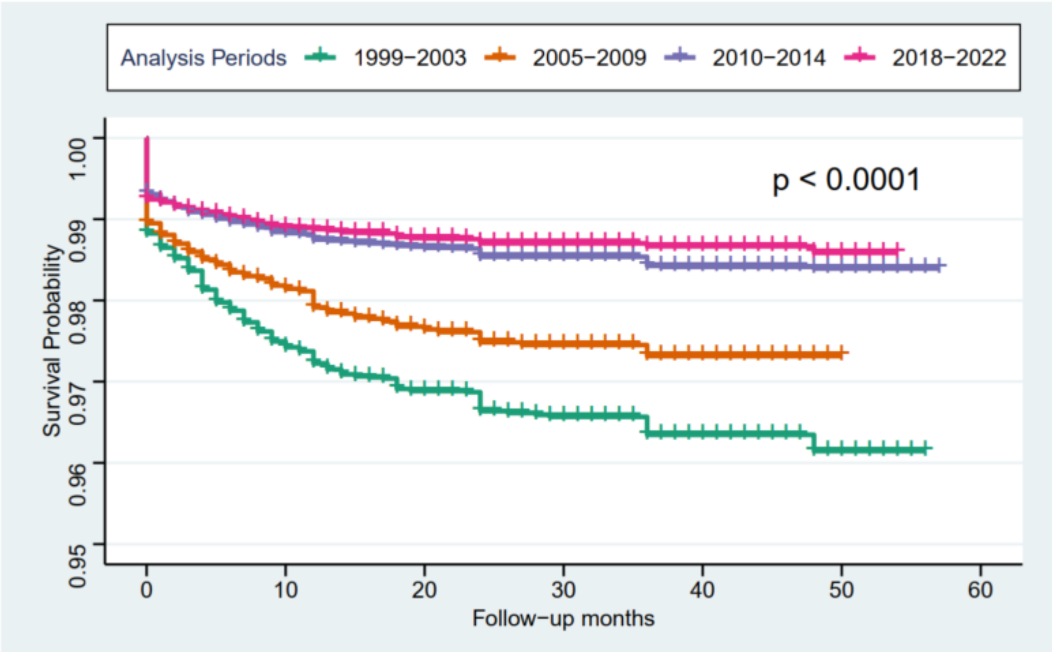

Figure 4.1: Survival Curve and Log-rank p-value

The Research Question

The analysis investigates whether survival distributions differ between different periods (1999-2003, 2005-2009, 2010-2014, and 2018-2022).

Define Hypotheses for the Log-Rank Test

The null hypothesis is that there are no differences in survival experience between the analysis periods while the alternative hypothesis is that at least one analysis period has a survival experience that differs from the others.

Description of the Survival Curves

Interpretation:

Overall Trend:

- The survival probability decreases over time across all periods.
- More recent periods (2010–2014 and 2018–2022) show higher survival probabilities than earlier periods (1999–2003 and 2005–2009).

Comparison Between Periods:

- The 1999–2003 cohort (green) has the lowest survival probability throughout the follow-up period.
- The 2005–2009 cohort (orange) performs slightly better but still has lower survival compared to later periods.
- The 2010–2014 cohort (purple) shows further improvement in survival.
- The 2018–2022 cohort (pink) has the highest survival probability, indicating significant progress in survival outcomes over time.

Statistical Significance:

The p-value is < 0.0001, indicating a statistically significant difference in survival probability across the different periods.

4.4 Survival Experience stratified by region

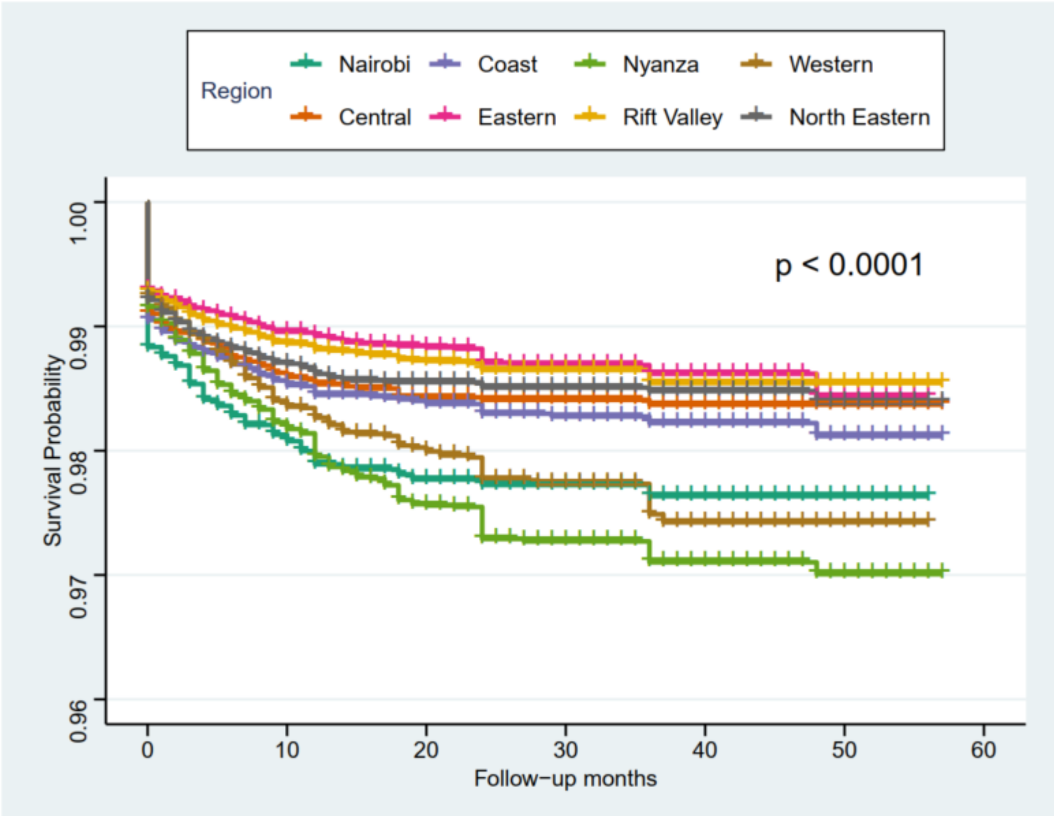

Figure 4.2: Survival Curve and Log-rank Tests

The Research Question

The objective of this analysis is to examine whether child survival probabilities differ significantly across different regions.

Define Hypotheses for the Log-Rank Test

The null hypothesis is that there is no difference in survival experience between the regions, while the alternative hypothesis is that at least one region has a survival experience that significantly differs from the others.

Description of the Survival Curves

Key Observations:

Overall Trend:

- Survival probability declines over time for all regions.
- There is a clear separation between regions, indicating differences in survival outcomes.
- The p-value (< 0.0001) suggests that these differences are statistically significant.

Regional Differences in Survival:

- Nyanza (green) and Western (brown) show the steepest decline in survival probability, indicating poorer survival outcomes.
- Coast (purple), Central (orange), and Nairobi (blue) have survival probabilities in the middle range.
- Eastern (pink), Rift Valley (yellow), and North Eastern (black) have the highest survival probabilities over time, suggesting better survival outcomes.

4.5 Survival Experience stratified by mother's level of education

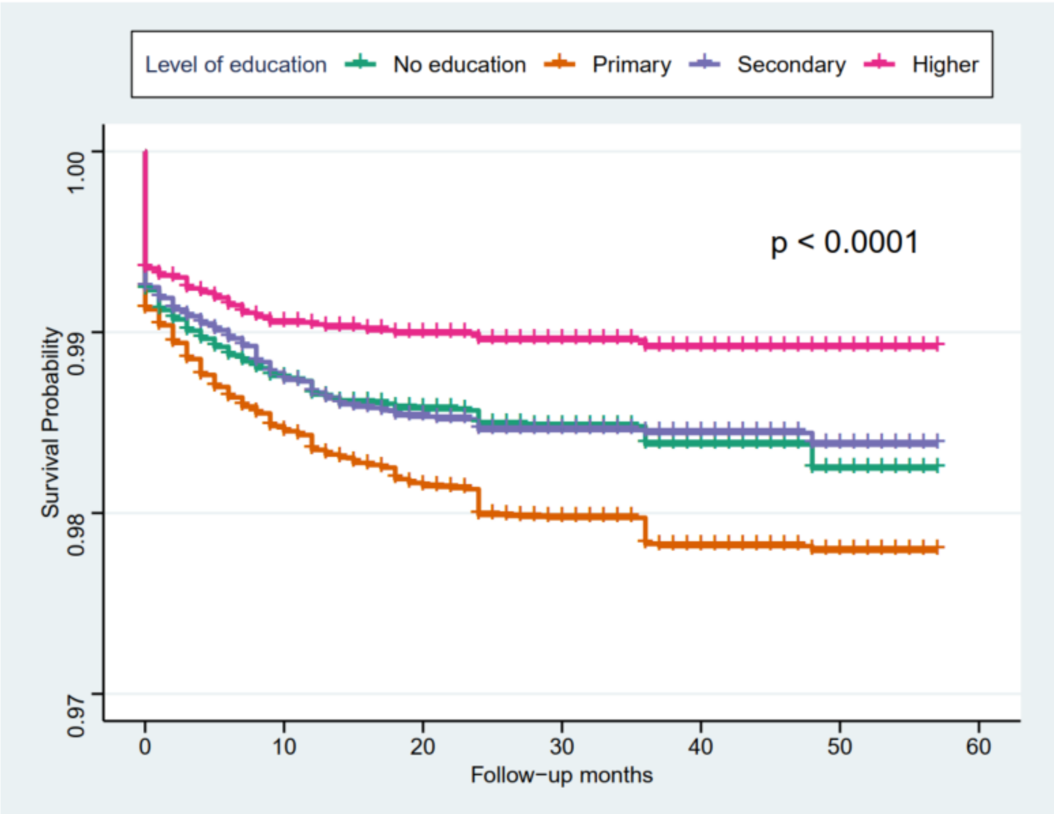

Figure 4.3: Survival Curve and Log-rank p-value

The Research Question

The objective of this analysis is to examine whether child survival probabilities differ significantly across different levels of education for the mother.

Define Hypotheses for the Log-Rank Test

The null hypothesis is that there is no difference in survival experience among children belonging to mothers with different levels of education, while the alternative hypothesis is that at least one group of children belonging to mothers with a particular level of education has a survival experience that significantly differs from the other levels.

Key Observations:

Overall Trend:

- Survival probability decreases over time across all education levels.
- There is a clear stratification, with higher education levels associated with better survival outcomes.
- The p-value (< 0.0001) indicates that the differences in survival between education levels are statistically significant.

Education Level and Survival:

- Children belonging to mothers with higher education (pink) have the highest survival probability throughout the follow-up period.
- Children belonging to mothers with secondary education (purple) and no education (green) have similar survival curves, though secondary education performs slightly better.
- Children belonging to mothers with primary education (orange) show the most pronounced decline in survival over time.

4.6 Survival Experience stratified by sex of the child

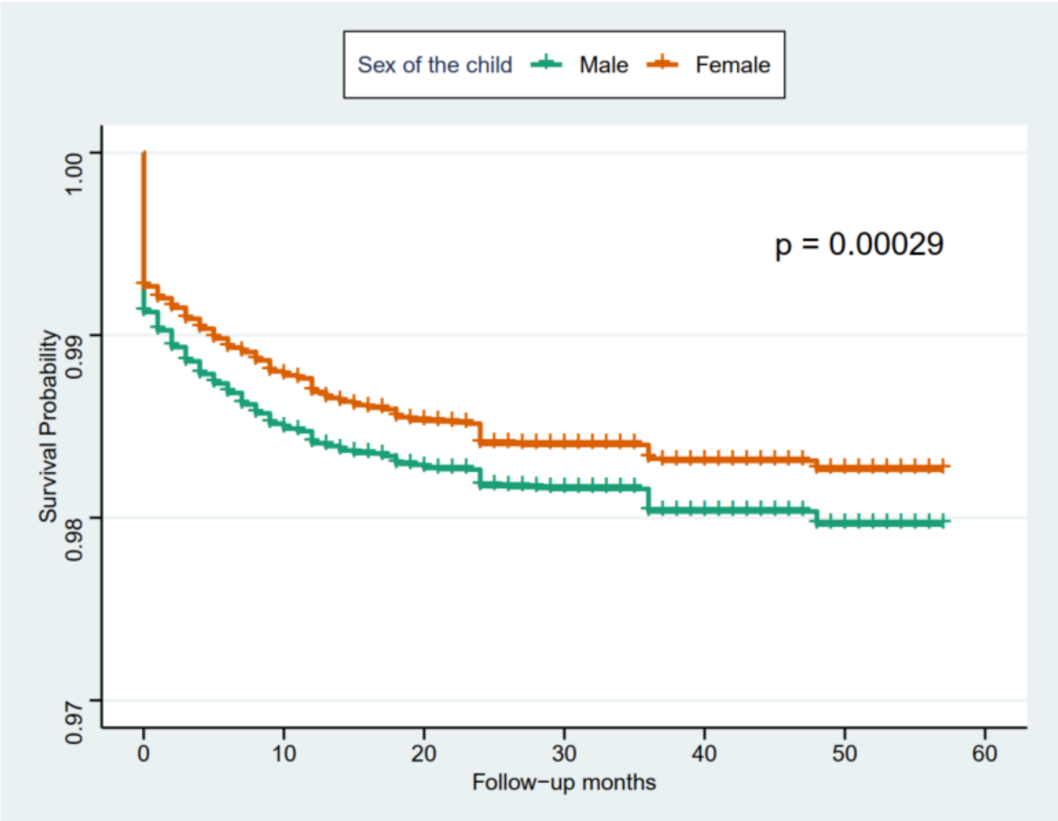

Figure 4.4: Survival Curve and Log-rank p-value

The Research Question

The objective of this analysis is to examine whether the survival experience differs based on the gender of the child.

Define Hypotheses for the Log-Rank Test

The null hypothesis is that there is no difference in survival experience based on gender while the alternative hypothesis is that the survival experience is different based on gender.

This Kaplan-Meier survival curve compares survival probabilities over time between male (green) and female (orange) children. The p-value of 0.00029 suggests that the survival differences between sexes are statistically significant.

Key Observations:

Overall Survival Trend:

- The survival probability decreases over time for both sexes.
- Female children (orange) have consistently higher survival probabilities compared to male children (green) throughout the follow-up period.

Sex-Based Survival Disparity:

- Female children show better survival outcomes at every time point.
- Male children have a steeper decline in survival, indicating a higher risk of mortality.
- The difference in survival remains consistent over time.

Statistical Significance (p = 0.00029):

- The p-value suggests that the observed differences are unlikely to be due to random variation.
- This indicates a real effect of sex on child survival.

4.7 Distribution of mortality rates across periods and regions

4.7.1 Distribution of mortality rates across periods

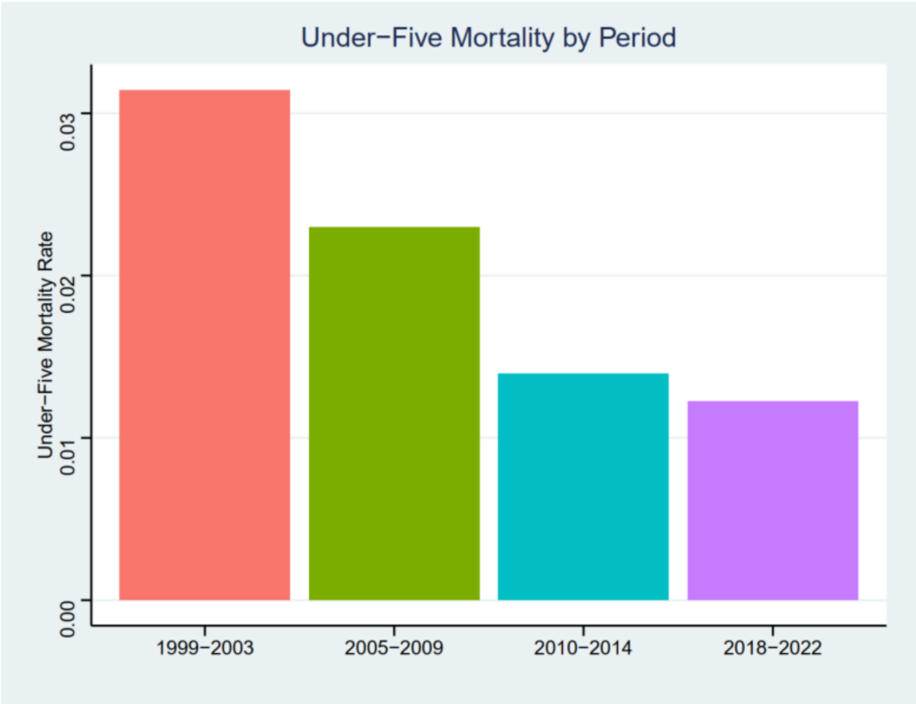

Table 1: Under-five mortality by period

	period	u5m	se
1999-2003	1999-2003	0.0335991	0.0018365
2005-2009	2005-2009	0.0235304	0.0018810
2010-2014	2010-2014	0.0144989	0.0007320
2018-2022	2018-2022	0.0115984	0.0006474

Figure 4.5: Mortality rates across Periods

Figure 4.5 and accompanying table present trends in mortality rates of those under five years of age in four time periods: **1999-2003, 2005-2009, 2010-2014, and 2018-2022.** The under-five mortality rate (u5m) represents the probability of a child dying before reaching the age of five, while the standard error (se) quantifies the uncertainty around the mortality estimates.

Key Observations

1. Declining Trend in Under-Five Mortality

- The mortality rate has consistently declined over time, from 0.033 (or 33 deaths per 1,000 live births) in 1999–2003 to 0.0116 (or 11.6 deaths per 1,000 live births) in 2018–2022.

- This represents a gradual but substantial reduction of nearly 22% over two decades, reflecting improvements in child survival.

2. Largest Drop Occurred Between 1999–2003 and 2005–2009

- The under-five mortality rate dropped from 0.0336 to 0.0235, a relative decrease of approximately 10%.

3. Slower Decline in Recent Years

- Between 2010–2014 and 2018–2022, the reduction in under-five mortality was less pronounced (from 0.0145 to 0.012).

- This suggests that while progress continues, achieving further reductions may require **targeted interventions** addressing specific causes of child mortality.

4.7.2 Distribution of mortality rates across regions

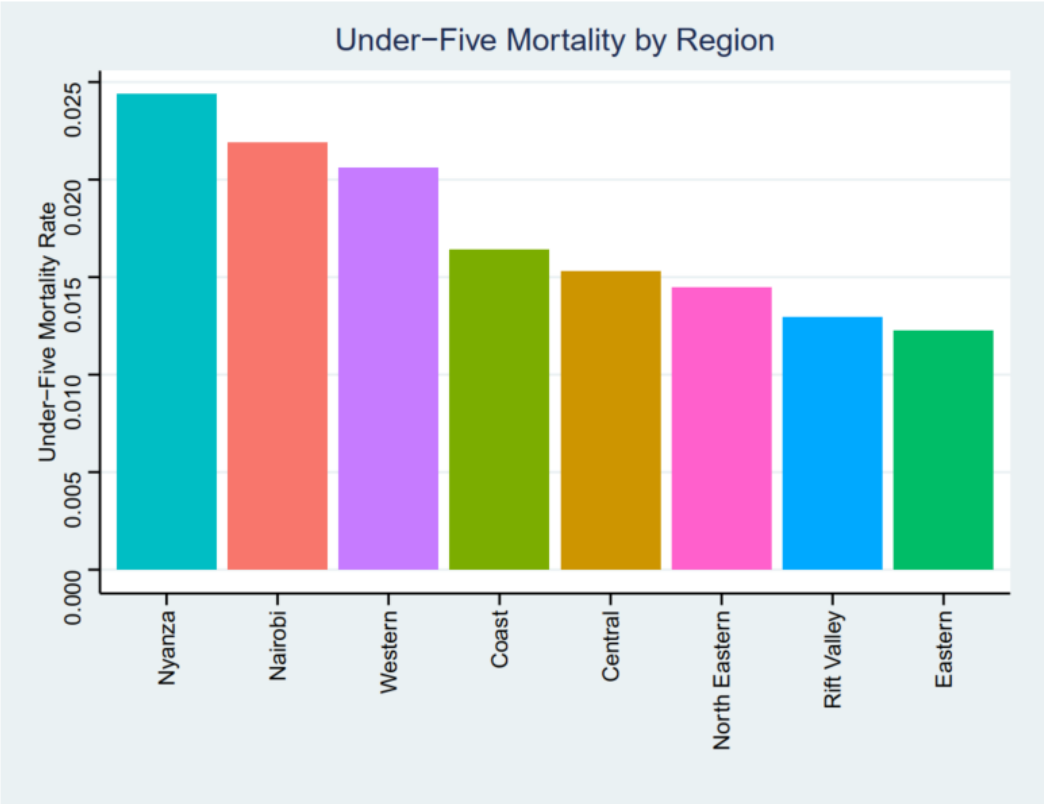

Table 2: Under-five mortality by region

	region	u5m	se
Nairobi	Nairobi	0.0202769	0.0027486
Central	Central	0.0131560	0.0013016
Coast	Coast	0.0167918	0.0014267
Eastern	Eastern	0.0140185	0.0012690
Nyanza	Nyanza	0.0275774	0.0015989
Rift Valley	Rift Valley	0.0136017	0.0007869
Western	Western	0.0191450	0.0015399
North Eastern	North Eastern	0.0149772	0.0014206

Figure 4.6: Mortality rates across Regions

Figure 4.6 and the corresponding table display the under-five mortality rate (u5m) by region. The mortality rate for children under five years reflects the probability that a child will die before age five. The table provides precise mortality rate estimates along with their standard errors (se), which indicate the level of uncertainty in each estimate.

Key Observations

1. Regional Disparities in Under-Five Mortality

- The highest under-five mortality rate is observed in Nyanza (0.0276 or 27.6 deaths per 1,000 live births), followed by Nairobi (0.0203 or 20.3 per 1,000).
- The lowest mortality rates are in the central (0.0132 or 13.2 per 1,000) and Rift Valley (0.0136 or 13.6 per 1,000).

Conclusion

There are **significant regional disparities in under-five mortality rates**, with **Nyanza and Nairobi facing the highest child mortality burdens**, while **Central and Rift Valley have the lowest**.

4.8 Estimating a model to determine the association of various factors with the hazard function

In this section, we present the results of Cox regression modeling. This helps determine significant factors associated with child survival. One of the key assumptions of the Cox model is that of proportional hazards, i.e. the hazard functions of different individuals should not cross over time. To test this assumption, we can either use graphical methods (K-M curves and log-log survival curves) or Schoenfeld residuals. In this study, we chose Schoenfeld residuals over graphical methods because they tend to be robust in many applications. A p-value of more than .05 implies the assumption has been met otherwise it's not met.

4.8.1 Unadjusted Cox Regression models

We start with simple single predictor models before running a multiple Cox regression model.

Model 1: Association of analysis periods with mortality

Table 4.2: Results for model 1

Characteristic	HR	95% CI	p-value
period			
1999–2003	–	–	–
2005–2009	0.74	(0.64, 0.87)	<0.001
2010–2014	0.44	(0.39, 0.49)	<0.001
2018–2022	0.39	(0.34, 0.44)	<0.001

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Testing the PH assumption

The results of testing the PH assumption using the Schoenfeld residuals are presented below;

Table 4.3: PH assumptions for model 1

	Chisq	df	p
period	55.1	3	<0.001
Global	55.1	3	<0.001

Since both the global and the individual p-values for the analysis period are less than .05, we conclude that the PH assumption has been violated.

Model 2: Association between regions and the hazard function

Table 4.4: Results for model 2

Characteristic	HR	95% CI	p-value
region			
Nairobi	–	–	–
Central	0.69	(0.53, 0.90)	0.007
Coast	0.75	(0.59, 0.95)	0.019
Eastern	0.56	(0.43, 0.71)	<0.001
Nyanza	1.11	(0.88, 1.40)	0.4
Rift Valley	0.59	(0.47, 0.74)	<0.001
Western	0.94	(0.74, 1.20)	0.6
North Eastern	0.65	(0.51, 0.84)	0.001

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Testing the PH assumption

Table 4.5: PH assumptions for model 2

	Chisq	df	p
region	88.2	7	<0.001
Global	88.2	7	<0.001

Both p-values are less than .05 implying that the PH assumption has been violated for this model as well.

Model 3: Association between sex of the child and the hazard function

Table 4.6: Results for model 3

Characteristic	HR	95% CI	p-value
Sex of the child			
Male	–	–	–
Female	0.85	(0.77, 0.93)	<0.001

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Testing the PH assumption

Table 4.7: PH assumptions for model 3

	Chisq	df	p
Sex of the child	1	1	0.3
Global	1	1	0.3

The p-values are all greater than .05, indicating that this model meets the PH assumption and can, therefore, explain variation in the survival experience.

Interpretation of the model

Reference Category (Male)

- The HR for males is not reported (denoted by "—"), as this group serves as the baseline for comparison.

Female (HR = 0.85, 95% CI: 0.77–0.93, p < 0.001)

- The hazard ratio of 0.85 indicates that female children have a 15% lower risk of the event (e.g., mortality) compared to male children.
- The confidence interval (0.77–0.93) does not include 1, meaning the result is statistically significant.
- The p-value (<0.001) confirms that the difference in risk between male and female children is highly significant.

Model 4: Association between wealth index and hazard function

Table 4.8: Results for model 4

Characteristic	HR	95% CI	p-value
Wealth Index			
Poorest	–	–	–
Poorer	1.11	(0.98, 1.27)	0.10
Middle	1.06	(0.93, 1.21)	0.4
Richer	0.97	(0.85, 1.12)	0.7
Richest	1.09	(0.95, 1.25)	0.2

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Testing the PH assumption

Table 4.9: PH assumptions for model 4

	Chisq	df	p
Wealth Index	4.2	4	0.4
Global	4.2	4	0.4

The p-values for the PH assumption test are greater than .05 indicating that the PH assumption is satisfied for this particular model.

Interpretation of the model

Reference Category (Poorest)

- The poorest category is the baseline (HR not reported).
- All other wealth categories are compared to this group.

Poorer (HR = 1.11, 95% CI: 0.98–1.27, p = 0.10)

- The HR of 1.11 suggests that individuals in the "poorer" category have an 11% higher risk of the event (e.g., mortality) compared to the poorest.
- However, the confidence interval (0.98–1.27) includes 1, meaning this result is not statistically significant.

- The p-value (0.10) indicates weak evidence against the null hypothesis.

Middle (HR = 1.06, 95% CI: 0.93–1.21, p = 0.40)

- HR of 1.06 suggests a 6% higher risk compared to the poorest.
- The CI (0.93–1.21) includes 1, indicating no significant difference.
- p-value (0.4) suggests no statistical significance.

Richer (HR = 0.97, 95% CI: 0.85–1.12, p = 0.70)

- HR of 0.97 suggests a 3% lower risk compared to the poorest.
- The CI (0.85–1.12) includes 1, meaning the difference is not statistically significant.
- p-value (0.7) suggests no significant effect.

Richest (HR = 1.09, 95% CI: 0.95–1.25, p = 0.20)

- HR of 1.09 suggests a 9% higher risk than the poorest.
- The CI (0.95–1.25) includes 1, indicating no significant difference.
- p-value (0.2) suggests no statistical significance.

Model 5: Association between smoking status and child mortality

Table 4.10: Results for model 5

Characteristic	**HR**	**95% CI**	**p-value**
Smoking Status			
No	–	–	–
Yes	1.21	(0.58, 2.54)	0.6

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Testing the PH assumption

Table 4.11: PH assumptions for model 5

	Chisq	df	p
Smoking Status	0.5	1	0.5
Global	0.5	1	0.5

The p-values for the PH assumption test are larger than .05 indicating that the PH assumption is satisfied for this model.

Interpretation of the model

Reference Category (Non-Smokers)

- The HR is not reported for non-smokers since it serves as the baseline group.

Smokers (HR = 1.21, 95% CI: 0.58–2.54, p = 0.6)

- The HR of 1.21 suggests that children belonging to mothers who smoke have a 21% higher risk of mortality compared to children belonging to non-smoking mothers.
- However, the confidence interval (0.58–2.54) includes 1, meaning this result is not statistically significant.

4.8.2 Adjusted Cox Regression model

We present the results from a Cox model that has been adjusted for multiple factors simultaneously. We will only consider independent variables that satisfy the proportional hazards assumption and discard those that fail to meet the assumption. Out of the independent variables that we had only wealth index, age of the mother, sex of the child, smoking status, type of residence, and preceding birth interval in months met the PH assumption. So the final model only considered these variables.

Table 4.12: Results for the adjusted model

Characteristic	HR	95% CI	p-value
Wealth Index			
Poorest	–	–	–
Poorer	1.13	(0.95, 1.35)	0.2
Middle	1.09	(0.91, 1.32)	0.4
Richer	0.99	(0.80, 1.22)	>0.9
Richest	1.41	(1.12, 1.79)	0.004
Age of the mother	1.00	(0.99, 1.01)	0.8
Sex of the child			
Male	–	–	–
Female	0.81	(0.72, 0.92)	<0.001
Smoking Status			
No	–	–	–
Yes	1.21	(0.54, 2.74)	0.6
Type of residence			
Urban	–	–	–
Rural	1.13	(0.95, 1.35)	0.2
Preced-birth-interval-months	1.00	(1.00, 1.00)	0.070

Abbreviations: CI = Confidence Interval, HR = Hazard Ratio

Interpretation of the model

Wealth Index

Being in the richest category is associated with a 41% higher hazard compared to the poorest group, and this is statistically significant ($p = 0.004$). Other wealth categories do not show significant differences.

Age of the mother The age of the mother does not significantly impact survival in this model.

Sex of the child

Female children have significantly better survival compared to males ($HR = 0.81$, $p < 0.001$).

Smoking status

Smoking status does not significantly impact survival in this model ($p = 0.6$).

Type of residence

There are no significant differences in survival between rural and urban residents.

Preceding birth interval in months

The preceding birth interval has a borderline effect (p = 0.070), meaning it may have a small impact but is not strongly significant.

4.9 Diagnostic tests

The Cox model has some underlying assumptions that are worth checking before using the model results with confidence. In this section, we will look at three assumptions of the Cox model. To conduct diagnostic checks we will use the residual methods. Three types of residuals can help us in this endeavor, these are martingale residuals, deviance residuals, and Schoenfeld residuals.

The martingale residuals are used to assess the non-linearity of the continuous predictors, deviance residuals which are symmetric transformations of the martingale residuals are used to examine influential observations, while the Schoenfeld residuals are used to check for the proportional hazards assumption.

4.9.1 Assessing the proportional hazards assumption

Table 4.13: PH assumptions for the adjusted model

	Chisq	df	p
Wealth Index	5.7	4	0.2
Age of the mother	2.1	1	0.2
Sex of the child	0.1	1	0.7
Smoking Status	1.0	1	0.3
Type of residence	0.0	1	0.8
Preceding-birth-interval	0.1	1	0.8
Global	10.9	9	0.3

The p-values are greater than .05 for both the predictors and the global test, therefore the proportional hazards assumption is not violated for this model.

Chapter 5

Discussions, Conclusions and Recommendations

5.1 Introduction

This chapter provides a synthesis of the research findings about the research objectives outlined in the Methodology chapter. It discusses how the research questions have been addressed, interprets the significance of the findings, presents conclusions, and offers recommendations for policy, practice, and future research.

The study set out to examine the trends and disparities in under-five mortality rates across different regions of Kenya over the period from 2003 to 2022. Recognizing the persistent public health challenge of child mortality, the research aimed to identify the socio-economic, demographic, and healthcare determinants that contribute to variations in child survival rates. By reconstructing data from the Kenya Demographic and Health Surveys (KDHS), the study employed survival analysis techniques, including Kaplan-Meier estimation and Cox proportional hazards regression, to analyze time-to-event data and determine significant factors influencing child mortality.

Specifically, the study sought to:

- Assess the extent of regional disparities in under-five mortality and examine variations across different time periods.
- Identify key determinants such as maternal education, household wealth, and child-specific factors that influence mortality rates.

The insights gained from this research are critical for policymakers and healthcare practitioners in designing targeted interventions to further reduce child mortality and address disparities between different regions of Kenya. The following sections discuss the findings in detail, their implications, and possible recommendations for action and future research.

5.2 Discussion

5.2.1 Addressing Research Questions

The study sought to investigate the extent of regional disparities in child mortality trends in Kenya and identify key socio-economic, demographic, and healthcare-related determinants. The specific research questions were:

1. What is the extent of regional disparities in under-five mortality trends in Kenya between 2003 and 2022?

2. What are the key socio-economic, demographic, and healthcare determinants influencing these regional disparities?

Based on the results:

- **Regional Disparities:** The study confirmed significant regional disparities in child mortality rates. Regions such as Nyanza and Western exhibited higher mortality risks compared to Nairobi and Central Kenya. The log-rank tests and Kaplan-Meier survival estimates demonstrated statistically significant differences in survival experiences across regions (p-value < 0.0001).

- **Key Determinants:** The adjusted Cox regression model identified maternal education, wealth index, and sex of the child as key determinants. Female children exhibited a lower hazard ratio ($HR = 0.81$, $p < 0.001$) compared to males. The wealthiest households had significantly higher hazards, contrary to expectations, suggesting potential biases in healthcare access or reporting issues.

- **Trends Over Time:** Child survival has improved over time, with later periods (2010–2014 and 2018–2022) showing significantly higher survival probabilities than earlier periods (1999–2003). The decreasing hazard ratios over time indicate the effectiveness of public health interventions and improved healthcare services.

5.2.2 Interpretation of Findings

These findings align with existing literature that highlights disparities in healthcare access, maternal education, and economic inequalities as primary drivers of child mortality variations. The survival improvement over time suggests that national and global initiatives, including immunization programs and maternal health interventions, have contributed to reduced under-five mortality.

However, the unexpected increase in hazard for the wealthiest households warrants further investigation. It may indicate under-utilization of healthcare services, differences in reporting child deaths, or regional disparities in healthcare access despite economic standing.

5.2.3 Limitations of the Study

While the study provides valuable insights, certain limitations should be acknowledged:

- **Data Quality:** The study relied on secondary data, which may be subject to reporting biases and misclassification errors.
- **Unmeasured Confounders:** Some potential determinants of child mortality, such as environmental factors and healthcare infrastructure quality, were not included in the KDHS datasets.
- **Censoring Issues:** Survival analysis methods require assumptions about censoring, which may affect the interpretation of results.

5.3 Conclusions

The study concludes that regional disparities in child mortality persist in Kenya, with significant differences in survival rates between regions. Maternal education, household wealth index, and the sex of the child are key determinants influencing child survival outcomes. Over the years, mortality rates have declined, suggesting that public health interventions have had a positive impact. As a country we are on track for achieving the 2030 target for reducing mortality rates to 25 deaths per 1000 live births, because according to the results as of the 2022 survey, the mortality rate was 11.6 or approximately 12 deaths per 1000 live births.

Despite improvements, disparities remain a challenge, and targeted interventions are necessary to address healthcare access inequalities and socio-economic disparities.

5.4 Recommendations

5.4.1 Policy Recommendations

To Strengthen maternal education programs to empower women with knowledge on child health and nutrition. Expand healthcare access in high-mortality regions, particularly Nyanza and Western Kenya, through targeted investments and, to improve data collection and reporting mechanisms to address inconsistencies in mortality estimates across different economic groups.

5.4.2 Recommendations for Future Research

To Conduct qualitative studies to understand healthcare-seeking behaviors among different socio-economic groups, explore the role of environmental factors and healthcare infrastructure in child mortality disparities, and apply machine learning techniques to enhance predictive modeling of child mortality trends.

By implementing these recommendations, policymakers and researchers can work towards achieving equitable child survival outcomes and further reducing under-five mortality in Kenya.

References

Adetunji, J. A. (1995). Infant mortality and mother's education in ondo state, nigeria. *Social Science & Medicine*, 40(2):253–263.

Anderson, R. N. and Rosenberg, H. M. (1998). Age standardization of death rates; implementation of the year 2000 standard.

Bell, R., Taylor, S., and Marmot, M. (2010). Global health governance: commission on social determinants of health and the imperative for change. *Journal of Law, Medicine & Ethics*, 38(3):470–485.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Breslow, N. E., Day, N. E., and Heseltine, E. (1980). Statistical methods in cancer research.

Chiang, C. (1984). The life table and its applications.

Cleves, M. (2008). *An introduction to survival analysis using Stata*. Stata press.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Croft, T. N., Marshall, A. M., Allen, C. K., Arnold, F., Assaf, S., Balian, S., et al. (2018). Guide to dhs statistics. *Rockville: ICF*, 645:292–303.

Doll, R. and Peto, R. (1981). The causes of cancer: quantitative estimates of avoidable risks of cancer in the united states today. *JNCI: Journal of the National Cancer Institute*, 66(6):1192–1308.

Ettarh, R. and Kimani, J. (2012). Determinants of under-five mortality in rural and urban kenya. *Rural and Remote Health*, 12(1):3–11.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Hill, C. L., Black, R. J., Nossent, J. C., Ruediger, C., Nguyen, L., Ninan, J. V., and Lester, S. (2017). Risk of mortality in patients with giant cell arteritis: a systematic review and meta-analysis. In *Seminars in arthritis and rheumatism*, volume 46, pages 513–519. Elsevier.

Imbo, A. E., Mbuthia, E. K., and Ngotho, D. N. (2021). Determinants of neonatal mortality in kenya: evidence from the kenya demographic and health survey 2014. *International Journal of Maternal and Child Health and AIDS*, 10(2):287.

Journals, B. (2020). Accelerating kenya's progress to 2030: understanding the determinants of under-five mortality from 1990 to 2015. *BMJ Journals2*.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481.

Keats, E. C., Macharia, W., Singh, N. S., Akseer, N., Ravishankar, N., Ngugi, A. K., Rizvi, A., Khaemba, E. N., Tole, J., and Bhutta, Z. A. (2018). Accelerating kenya's progress to 2030: understanding the determinants of under-five mortality from 1990 to 2015. *BMJ global health*, 3(3):e000655.

Kedogo, J. (2023). Institute of economic affairs. reducing child mortality in kenya is both urgent and possible. institute of economic affairs. https://ieakenya.or.ke/blog/reducing-child-mortality-in-kenya-is-both-urgent-and-possible/. *health*.

Kim, H.-J., Fay, M. P., Feuer, E. J., and Midthune, D. N. (2000). Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in medicine*, 19(3):335–351.

Kimani-Murage, E. W., Fotso, J.-C., Egondi, T., Abuya, B., Elungata, P., Ziraba, A. K., Kabiru, C. W., and Madise, N. (2014). Trends in childhood mortality in kenya: the urban advantage has seemingly been wiped out. *Health & place*, 29:95–103.

Macharia, P. M., Giorgi, E., Thuranira, P. N., Joseph, N. K., Sartorius, B., Snow, R. W., and Okiro, E. A. (2019). Sub national variation and inequalities in under-five mortality in kenya since 1965. *BMC public health*, 19:1–12.

Macharia, P. M., Joseph, N. K., Sartorius, B., Snow, R. W., and Okiro, E. A. (2021). Subnational estimates of factors associated with under-five mortality in kenya: a spatio-temporal analysis, 1993–2014. *BMJ global health*, 6(4):e004544.

Misselhorn, M. and Harttgen, K. (2006). A multilevel approach to explain child mortality and undernutrition in south asia and sub-saharan africa.

Nasejje, J. B., Mwambi, H. G., and Achia, T. N. (2015). 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*, 15:1–12.

Preston, S. (2000). Demography: Measuring and modeling population processes. *(No Title)*.

Rutstein, S. O., Rojas, G., et al. (2006). Guide to dhs statistics. *Calverton, MD: ORC Macro*, 38:78.

Tesema, G. A., Teshale, A. B., and Tessema, Z. T. (2021). Incidence and predictors of under-five mortality in east africa using multilevel weibull regression modeling. *Archives of Public Health*, 79:1–13.

Victora, C. G., Wagstaff, A., Schellenberg, J. A., Gwatkin, D., Claeson, M., and Habicht, J.-P. (2003). Applying an equity lens to child health and mortality: more of the same is not enough. *The Lancet*, 362(9379):233–241.

Appendix A

R code Repository

A.1 GitHub Repository

Here is the link to my GitHub repository that contains the R scripts used in conducting data analysis for this master thesis. https://github.com/rikoprogrammer/Msc_THESIS_codes

Appendix B

Ethical approval

Strathmore
UNIVERSITY

26th February 2025

Mr Shikuku Eric,
eric.shikuku@strathmore.edu

Dear Mr Shikuku,

<u>RE: Trends in the Under-five Mortality and Disparities between Regions in Kenya</u>

This is to inform you that SU-ISERC has reviewed and **approved** your above **SU-masters** proposal. Your application reference number is **SU-ISERC2662/25.** The approval period is from **26th February 2025 to 25th February 2026.**

This approval is subject to compliance with the following requirements:

i. Only approved documents including (informed consents, study instruments, MTA) will be used.
ii. All changes including (amendments, deviations, and violations) are submitted for review and approval by SU-ISERC.
iii. Death and life-threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to SU-ISERC within 72 hours of notification.
iv. Any changes anticipated or otherwise that may increase the risks or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to SU-ISERC within 72 hours.
v. Clearance for the export of biological specimens must be obtained from relevant institutions.
vi. Submission of a request for renewal of approval at least 60 days prior to the expiry of the approval period. Attach a comprehensive progress report to support the renewal.
vii. Submission of an executive summary report within 90 days of completion of the study to SU-ISERC.

Before commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology, and Innovation (NACOSTI) https://research-portal.nacosti.go.ke/ and obtain other clearances needed.

Yours sincerely,

Mr Ambrose Rachier,
Chairperson: SU-ISERC

Appendix C

Similarity index

Eric Shikuku

Trends in the Under-five Mortality and Disparities between Regions in Kenya.pdf

Strathmore University (Main Account)

Document Details

Submission ID

trn:oid:::2945:275119342

Submission Date

Mar 28, 2025, 1:35 PM GMT+3

Download Date

Mar 28, 2025, 2:38 PM GMT+3

File Name

Trends in the Under-five Mortality and Disparities between Regions in Kenya.pdf

File Size

1.2 MB

54 Pages

9,048 Words

53,089 Characters

turnitin Page 1 of 66 - Cover Page Submission ID trn:oid:::2945:275119342

25% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

- **183** Not Cited or Quoted 23%
  Matches with neither in-text citation nor quotation marks
- **24** Missing Quotations 3%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

18% Internet sources
13% Publications
21% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.