Survival Analysis of Child Mortality in Kenya: An Analysis of 2014 Kenya Demographic and Health Survey Data

Elda Naliaka Watulo*, Anthony Kibera Wanjoya

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

Email address:
elda.watulo@gmail.com (E. N. Watulo), awanjoya@gmail.com (A. K. Wanjoya)

*Corresponding author

To cite this article: Elda Naliaka Watulo, Anthony Kibera Wanjoya. Survival Analysis of Child Mortality in Kenya: An Analysis of 2014 Kenya Demographic and Health Survey Data. International Journal of Statistical Distributions and Applications. Vol. 7, No. 2, 2021, pp. 57-71.
doi: 10.11648/j.ijsd.20210702.14

Received: June 11, 2021; Accepted: June 23, 2021; Published: June 28, 2021

Abstract: Kenya is one of the Sub-Saharan African countries that experience the highest risk of children dying before reaching the age of five years. The effects of demographic, environmental, and socio-economic factors play a significant role in under-five mortality (U5CM). The study aimed to determine the risk factors of under-5 child mortality in Kenya. In an attempt to determine the risk factors of under-5 mortality, survival analysis using both Kaplan-Meier and Cox hazard methods of live births within the Kenyan population based on the 2014 Kenya Demographic and Health Survey (KDHS) data was adopted. All children born in the period between 2009 and 2014 (n=83,591) were included in the study. The outcome variable was the all-cause under-5 mortality. Within the observation period, a total of 6,123 child deaths were recorded. The under-5 mortality rate in Kenya is strongly associated with the mother's education, region, place of residence, preceding birth interval, birth order, the total number of children ever born, mother's occupation, and type of toilet facility. The results indicated that a child born in Nyanza is twice more likely to die than that born in the Central region of Kenya. Male children have a higher risk of dying before the age of five than their female counterparts. The risk of experiencing U5CM increased among children born in rural areas compared to those born in urban areas. In summary, the study suggests that under-five child mortality is still a problem in Kenya. The government and implementing partners should allocate resources towards maternal and child healthcare, and implement interventions on women empowerment, and scale up health education among mothers. Both national and county governments should allocate resources to ensure access and use of modern contraceptives to improve child spacing. There is a need for the central government to implement socio-economic development interventions that reduce regional disparities.

Keywords: Under-Five Child Mortality, Survival Models, Cox-proportional Hazard, Kaplan-Meier, Kenya Demographic and Health Survey

1. Introduction

Child mortality is an indicator of the overall health and well-being of a nation. Not only does it reflect the magnitude of the health issues that are directly responsible for the deaths of children, such as malnutrition, respiratory, and diarrheal infections, but also the net effect of other factors such as pre-and postnatal care of mothers and children, as well as the environmental conditions in which the child is exposed.

A report from the United Nations Inter-Agency Group for Child Mortality Estimation (UN-IGME) [1] indicates that globally, the under-five mortality rate fell to 53 deaths per 1,000 live births in 2009 from 78 deaths per 1,000 live births in 1999; the highest proportion of these deaths occurred in developing countries. The report further indicated that most of the deaths occurred in Sub-Saharan African countries. The most significant reduction (of over 50%) in the under-five child deaths was, for example, witnessed in Latin America and the Caribbean, Northern Africa, Eastern Asia, Western Asia, and South-Eastern Asia. The lowest reduction was recorded in Sub-Saharan Africa (SSA), having posted a 39% drop. The disparities in regions are attributed to numerous factors, including differences in environmental, socioeconomic, and health conditions.

Kenya is among the SSA countries that are still experiencing the highest under-five child mortality. According to the 2013 UN-IGME report [1], one in every nine
children in SSA dies before attaining five years. From 1965 to 1980, Kenya enjoyed an impressive decline in the under-five mortality rate [2], which was twice the rate of an average Sub-Saharan country during the period. An analysis carried out by several research bodies in Kenya on the 1998 Kenya Demographic Health Survey Data indicated that the decline in U5MR had not just slowed down but had been reversed during the 1990s. Nevertheless, another analysis on the 2008/9 Kenya Demographic Health Survey data indicated a reversal in mortality rates during the late 1990s and early 2000s -the rate had declined by a whopping 36 percent (from 115 deaths to 74 deaths per 1000 live births). This remarkable decline in the under-five mortality levels was attributed to the government's efforts to improve health services [3].

One of the main objectives of the fourth Millennium Development Goals, MDG4, and the third Sustainable Development Goal, SDG 3, is to reduce the child mortality rate, which by implication, increases the child survival rate. Most of the factors that cause under-five deaths in Kenya today stem from problems long since solved in developed countries, such as poor sanitation, lack of access to clean water, and lack of protection against malaria-causing mosquitoes [1].

The purpose of this study is to examine the determinants of under-five child mortality in Kenya using the 2014 Kenya Demographic Health Survey data. The study's rationale is that it offers new information on the determinants and trends of child mortality as a basis for the planning and implementation of health policies, programs, and projects. This study is also relevant as very few studies have been done on this topic in Kenya, and none has been done using the latest health survey data, 2014 KDHS. These insights are crucial in making recommendations on possible focal points for child survival programs and interventions to curb the high child mortality in Kenya and other Sub-Saharan countries. In addition, this study could be used as a reference for dealing with future causes of under-five mortality. Therefore, the findings of this study should be of great use to the Ministry of Health and other stakeholders in the health sector in designing and implementing child health intervention projects and programs.

2. Related Work

Demographic researches carried out by Schultz [4] and by Mosley and Chen [5] classified factors affecting under-5 child mortality into social demographic, socio-economic, and environmental factors.

2.1. Social Demographic Factors

The sex of a child, birth order, and mother’s age at first birth are among the socio-demographic factors that influence child mortality, as pointed out by Mosley and Chen [5]. Other similar studies were carried out by Rutstein [6] and Davanzo et al. [7].

In Kenya, the mortality patterns by maternal age are usually U-shaped. Children born to relatively younger mothers (under the age of 20 years) and older (above the age of 40 years) mothers experience higher mortality rates than those born to mothers between the age of 20 and 40 years of age [8]. According to Kibet [8], younger mothers are inexperienced in taking care of their children. Besides, a child born to a younger mother tends to be malnourished, underweight, and in some cases, anemic, a combination of which increases mortality risk.

Birth intervals have been shown to have a strong effect on both infant and child mortality. Shortly spaced birth intervals are associated with adverse birth outcomes, higher morbidity rates during pregnancy, and higher infant and child mortality rates [9]. Kumar and File [10] used cross-tabulation techniques on Ethiopian DHS 2005 survey data to examine chosen bio-demographic, socioeconomic, and maternal health care predictors of child mortality in Ethiopia. The results indicated that birth interval is among the social demographic factors that significantly increase child mortality.

Mutunga [11] analyzed the Kenyan Demographic Health Survey 2003 data to examine the effect of environmental and social demographic predictors of infant and child mortality in Kenyan urban areas. The results indicated that child mortality was higher for children with birth intervals of less than two years. Birth order is strongly associated with child mortality in many studies, such as in the study by Mosley and Chen [5]. According to these studies, first-born children are more likely to die, maybe because of their mother’s age or birth complications. The higher risk of mortality to first births may be because such births occur mainly to younger women who are inexperienced in caring for infants.

Several studies have linked the sex of a child to child mortality, with boys experiencing higher chances of dying before the age of five years than their counterpart females. Experts have explained this by the sex differences in biological and genetic makeup, with male children being biologically weaker and more vulnerable to infections and early death.

Breastfeeding duration is also linked to child mortality. Breast milk is healthy and essential for a child’s growth as it is loaded with nutrients, antibodies, antioxidants, and hormones necessary for child survival. Children who are exclusively breastfed for the first six months of life and continue breastfeeding until at least two years of age have a stronger immune system than those who are not [12]. During birth, infants are born with strong protection against infections due to the antibodies passed to them from mothers through the placenta. After birth, breast milk continues supplying additional antibodies to keep the children strong and healthy. The difference narrows as the children are introduced to other foods.

2.2. Socioeconomic Factors

Socioeconomic factors are covariates that act through the proximate determinants to affect the under-five mortality rate. They include the place of residence, mother’s educational level, employment status of the mother, and region.

The employment status of the mother has an impact on the child mortality rate. The income a mother receives from her employment is a factor as it determines how she can provide for the basic needs of the family. Employed mothers have the financial ability to provide quality health care for their children, hence increasing survival prospects of the under-five children.
However, mothers who work away from home may not provide
the necessary child care, especially if the income is low.

Lack of proper breastfeeding for the first months after birth
lowers survival chances [13]. In this case, it is the mother’s
absence to care for her child rather than the employment that
negatively affects the survival of a child. The results may be
different for those mothers working from home.

Place of residence is also one of the determinants of child
health and mortality. According to Stalling [14], children
living in urban areas where safe drinking water, proper
sanitation, and better health care services are available,
experience lower mortality rates. According to Woldemicael
[15], place of residence affects a mother’s exposure to both
education and the extent to which proper sanitation, health
care facilities, and clean water are available, thus affecting the
survival chances of a child.

Education is one of the prerequisites to a country’s
socioeconomic development-an educated population tends to
be wealthier and healthier. According to Mutunga [16], higher
educational levels are usually associated with lower mortality
rates since education exposes mothers to knowledge about
better nutrition, childhood diseases and treatment, and the use
of family planning methods to space births. Goro [17] used a
multivariate logistic regression model to analyze Ghanaian data
from 1993, 1998, and 2003 DHS surveys to examine the factors
affecting child mortality in three northern regions. He found out
that a mother’s level of education, mother’s marital status, and
birth order are strong determinants of infant mortality.

Another research carried out in Kenya by Hill [2] indicated
that mother’s educational level plays a role in infant and child
mortality. According to these studies, literate women
understand infection and disease prevention measures such as
the use of mosquito nets and vaccines. They are more likely to
take sick children to the hospital as early as possible and
follow the doctors’ instructions. They understand germ theory
and put in place hygiene measures as household priorities.
Many studies show that family size influences rates of
morbidity and mortality. According to Woldemicael et al. [15],
when many people stay together, the risk of contracting
pathogens increases, and hygiene may deteriorate. A
household with many children stands a higher risk of having
disease infections due to crowding and competition for
resources and the mother’s attention [15].

2.3. Environmental Factors

The environment has a significant impact on the overall
health of society. Environmental factors that affect health
include drinking water sources, type of toilet facilities, and
waste disposal methods. Most of these factors are often
affected by place of residence and socioeconomic status
[6].

Several studies have outlined the health benefits of
consuming safe and clean water. Analyzing DHS data from
8 sub-Saharan African countries, Fayehun [18] found that
in countries with lower child mortality, such as Lesotho and
Namibia, the number of children living in households with
improved sources of water is higher than in households
with poor sources of drinking water. According to Mutunga
[16], despite accessing water from safe and improved
sources, fetching it using dirty containers as well as
improper storage can contaminate it with infection-causing
organisms.

Studies indicate that children staying in houses with
improved toilet facilities are less likely to fall sick than those
in households with no or unimproved toilet facilities.
Research conducted in Ghana shows that the risk of
contracting infections is significantly associated with toilet
facilities, with children in houses with toilets being about 50%
less likely to contract diseases than those living in houses with
no toilets [19].

3. Concepts of Survival Models

Survival analysis involves analyzing time-to-event
occurrences. The outcome of interest is the time to event,
generically referred to as survival time, failure time, or event
time, usually measured in days, weeks, months, or years.
Survival analysis methods analyze data that have three main
features:

i. The response or dependent variable is the waiting time
   until a well-defined event occurs
ii. The observations are censored
iii. There are predictors that we wish to assess or control the
    effect on the waiting time

3.1. Survival Analysis Data Structure

In survival analysis, there are no exact starting and ending
points. All the observations do not always start at zero. A
subject can enter the study at any time. All the subjects are
brought to a common starting point where the time \( t \) is zero.
At time zero, all subjects have the survival probabilities equal
to one, that is their chance of not experiencing the event of
interest, such as death, is 100%.

3.2. Censoring

Survival analysis the event of interest may not be observed
for some individuals during the study period; this is what we
refer to as censoring. Censoring is what makes the survival
analysis methods different from other methods of analysis.
Censoring may arise in one of the following three ways:

a) An individual has not yet experienced the event of
   interest, such as death at the end of the study
b) An individual is lost to follow-up in the course of the study
   period
   c) An individual experiences a different event other than
      the event of interest, making follow-up impossible

3.3. Important Functions in Survival Analysis

The first step in carrying out the analysis of survival data is
the estimation of the distribution of survival times.

3.3.1. Survival Function, \( S(t) \)

The survivor function or the survival distribution time
describes the survival experiences of the population of interest. Let \((T \geq 0)\) denote a positive random variable that represents the time to the event, and \(F(t) = Pr(T \leq t)\) denote a Lifetime Distribution Function, which is the probability that an event of interest will happen before time \(t\). The survival experience of individuals under study is described by the survival function, \(S(t)\).

Given the time to event \(T\), survival function, \(S(t)\) is given by:

\[
S(t) = P[r(T > t)] = 1 - F(t)
\]

It gives the probability that an individual will survive from time of origin to a specified future time \(t\). The survival function is between zero and one (inclusive), is a non-increasing function of \(t\), and it approaches zero as times goes to infinity.

### 3.3.2. The Hazard Function (\(\lambda\))

Also referred to as the instantaneous death rate, instantaneous failure rate or the force of mortality, the hazard function provides one of the ways of modelling data distribution in survival analysis. The hazard function gives the probability that if a subject survives to time \(t\), it will succumb to the event in the next instant. In other words, it is the probability that the individual will experience an event of interest within a small time interval, as long as the individual has survived until the beginning of that interval.

It is given by:

\[
\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}
\]

**NOTE:** In contrast to the survival function, which focuses on not experiencing the event, the hazard function focuses on the event occurring.

### 3.3.3. The Cumulative Hazard Function, \(\Lambda(t)\)

This is a function that describes the accumulated risk up to time \(t\).

\[
\Lambda(t) = \int_0^t \lambda(u) du
\]

### 3.4. Distribution-Free Survival Methods

For non-parametric estimation of survival data, the survival and hazard functions are used to summarize the data. The aim of using non-parametric methods to analyze survival data is to come up with graphical summaries of the survival times for a given group of subjects in the study. These graphical summaries are for the survival and hazard and functions. For a more detailed analysis, the median and other percentiles can be obtained.

#### 3.4.1. The Empirical Survival Function

Assuming independence of observations and that no data point is censored, the survival function \(S(t)\) can be estimated using the empirical survival function. The empirical function estimates the survival function in the absence of censored data. It is given by:

\[
\hat{S}(t) = \frac{\text{Number of individuals with } T > t}{\text{Total sample size}}
\]

\(\hat{S}(t) = 1\) for all values of \(t\) before the first failure and \(\hat{S}(t) = 0\) after the final occurrence of the event of interest.

Note that \(F_n(t) = 1 - \hat{S}_n(t)\) is the empirical cumulative distribution function.

#### 3.4.2. The Kaplan-Meier Estimator

The Kaplan-Meier method is a widely used non-parametric estimator of the survival function. We can use the Kaplan-Meier method to obtain univariate descriptive statistics for the survival data and compare survival experiences for groups of subjects.

The Kaplan-Meier estimator or the product limit estimator provides an estimate of the survival function for censored observations. The Kaplan-Meier estimator of the survivorship function or the probability of survival, \(S(t) = P(T > t)\) is given by:

\[
\hat{S}(t) = \prod_{i=1}^{k} \left( 1 - \frac{d_i}{n_i} \right) = \prod_{i=1}^{k} \left( 1 - \frac{d_i}{n_i} \right)
\]

where:

- \(i\) is a set of \(K\) distinct uncensored failure times observed in the sample.
- \(d_i\) is the number of failures at \(i\)
- \(n_i\) is the number of individuals "at risk" right before the \(i^{th}\) failure time (everyone who died or censored at or after time \(i\))

The standard error is therefore given by:

\[
s.e \hat{S}(t) = \hat{S}(t) \sum_{i=1}^{K} \frac{d_i}{n_i(n_i - d_i)^{1/2}}
\]

#### 3.4.3. Cumulative Hazard Function

Given \(\hat{S}(t)\) as the Kaplan-Meier estimator for the survival function, then the estimator to the cumulative hazard function is given by:

\[
\hat{\Lambda}(t) = - \sum_{i=1}^{k} \ln \left( 1 - \frac{d_i}{n_i} \right)
\]

Using Taylor series of expansion, we get:

\[
\ln \left( 1 - \frac{d_i}{n_i} \right) = - \frac{d_i}{n_i} - \left( \frac{d_i}{n_i} \right)^2 + \ldots - \left( \frac{d_i}{n_i} \right)^k
\]

Ignoring higher order terms, the cumulative hazard function becomes:

\[
\hat{\Lambda}(t) = \sum_{i=1}^{k} \frac{d_i}{n_i}
\]

Furthermore, in case of censoring:
\[ n_i = n_{i-1} - d_{i-1} - c_{i-1} \]

where, \( c_i \) is the number of censored observations.

Note that:

a) \( \hat{S}(t) \) only changes at failure times
b) \( \hat{S}(t) \) is until the first failure time
c) \( \hat{S}(t) \) only goes to 0 if the last observation is not censored

The Kaplan-Meier estimator equals the empirical survival estimator when there is no censoring.

3.5. Parametric Survival Methods

These models assume that survival times follow a specific probability distribution. The most used distributions include Exponential, Weibull, Gama, and Log-Logistic distributions.

3.5.1. Exponential Distribution

The exponential model as a constant hazard rate \( \lambda(t) = \lambda \).

In other words, the probability that an individual will experience the event of interest within a small time interval given that the individual has not experienced the event at the beginning of the interval is constant for any time period.

Thus the survival function \( S(t) \), which is the probability that an individual survives beyond time \( t \) is:

\[ S(t) = \exp(-\lambda t) \]

And the density is:

\[ f(t) = \lambda \exp(-\lambda t) \]

The hazard function \( \lambda(t) \) can be derived from both the density and survival as follows:

\[ \lambda(t) = \frac{f(t)}{S(t)} = \lambda \]

Finally, the cumulative hazard function is given by:

\[ \Lambda(t) = \lambda t \]

Due to a constant hazard rate, the exponential distribution has a unique property called the loss of memory property.

3.5.2. Weibull Distribution

The Weibull distribution is a more general model with two parameters; the scale, \( \lambda \), and shape, \( \kappa \), parameters. Due to its flexibility, the Weibull model can be used to model survival times of populations whose hazard rate is assumed to be constant \( (\kappa = 1) \), increasing \( (\kappa > 1) \), and decreasing \( (\kappa < 1) \). When \( \kappa = 1 \), the Weibull distribution simplifies to an Exponential distribution.

The cumulative density function \( F(t) \) is given by:

\[ F(t) = 1 - \exp(-[t\lambda]^{\kappa}) \]

The survival function denoted by \( S(t) \) is given as:

\[ S(t) = \exp(-\lambda t^{\kappa}) \]

The density function \( f(t) \) is given by:

\[ f(t) = \frac{d}{dt} S(t) = \kappa \lambda t^{\kappa-1} \exp(-\lambda t^{\kappa}) \]
The hazard rate is given by:

\[ \lambda(t) = \kappa \lambda t^{\delta - 1} \]

It follows that the cumulative hazard function for the

\[ H(t) = \int_0^t \lambda(u) \, du \]

3.5.3. Log-normal Distribution

Just like the Weibull distribution, the Log-normal model is a very flexible distribution. The distribution can fit many types of failure times data. It also has two parameters for scale, \( \delta \), and shape, \( \tau \).

If \( T \) has log-normal distribution, then:

\[ Y = \log(T) = \delta + \tau W \]

Where \( W \) is the standard logistic distribution with a density given by:

\[ f_W(\tau) = \frac{e^\tau}{(1 + e^\tau)^2} \]

The CDF is given by:

\[ F_W(\tau) = \frac{e^\tau}{(1 + e^\tau)} \]

The survival function is the complement of the CDF given by:

\[ S_W(\tau) = \frac{1}{1 + e^\tau} \]

Therefore, for the variable \( T \), the survival function is given by:

\[ S_T(\tau) = \frac{1}{(1 + \lambda t)^p} \]

where \( p = \frac{1}{\delta} \) and \( \delta = -\log \lambda \)

When the observations follow a log-normal distribution, the log the observations follows a normal distribution with mean \( \mu \) and standard deviation \( \delta \). This property makes the log-normal distribution very flexible as the failure times data can easily be analyzed as normally distributed data.

3.6. Regression Models in Survival Analysis

Just like in linear and logistic regression, the statistical goal of survival analysis is to establish some measure of the effect that will describe the relationship between the time to failure and some covariates.
Several survival models are used to analyze the relationship between survival time and a set of explanatory variables. These methods include parametric, semi-parametric, and non-parametric approaches.

### 3.6.1. Parametric Regression Models

The most commonly used parametric regression models include exponential and Weibull distributions. Just like with logistic regression, parameter estimates in parametric survival models are obtained using maximum likelihood estimation.

Therefore, we can use the same procedures for testing and constructing confidence intervals in parametric survival analysis as we do for logistic regression.

#### Exponential Regression Model

The survival function of $T$ at covariate values $X = (x_1, x_2, ..., x_n)^T$ can be shown to be:

$$S(t|x) = \exp \left(-\left(t e^{-X^T \beta}\right)^\delta\right)$$

where $\beta = (\beta_0, \beta_1, ..., \beta_n)^T$ is a vector of regression coefficients.

The hazard function is given by:

$$\lambda(t|x) = \lambda \exp(X'\beta)$$

In log terms, the log hazard function is given by:

$$\log\lambda(t|X) = \left(\frac{1}{\delta} - 1\right) \log t - \log\delta - X^T \left(\frac{\beta}{\delta}\right)$$

The model transforms into a linear model for $Y = \ln(T)$.

#### Weibull Regression Model

The hazard function is normally given by:

$$h(t) = \delta \lambda^\delta t^{\delta - 1}$$

To include the covariate vector $X$, we write the hazard given $X$ as:

$$\lambda(t|X) = \lambda_0(t) \cdot \exp(X'\beta)$$

$$= \delta \lambda_0^\delta t^{\delta - 1} \cdot \exp(X'\beta) = \delta \lambda^\delta t^{\delta - 1}$$

Just like the exponential model, we get the following result in log-scale:

$$\log(\lambda(t|x)) = \log(\delta) + \delta \log(\lambda) + (\delta - 1) \log(t)$$

$$\log(t) = \log(\delta) + \delta \log(\lambda) + X'\beta + (\delta - 1) \log(t)$$

### 3.6.2. Non-parametric Regression Model: The Cox Proportional Hazard Model

Cox proportional hazard model (Cox PH) regression model is a popular semi-parametric model for analyzing survival data.

It investigates the relationship between time-to-event and the predictors through the use of the hazard function. The Cox model is based on the assumption that the predictors have an increasing effect on the hazard function and that the effect is constant over time.

The hazard function is given by:

$$\lambda(t|x) = h_0(t) e^{X'\beta} = \lambda_0(t) e^{X'\beta x_1}$$
where
\[ \lambda_0(t) \] is the baseline hazard, which does not depend on the covariates,
\[ \lambda(t|x) \] is the hazard at time t for an individual with a set of predictor variables \( X \)
\[ \beta \] are the parameters describing the effect of the predictor variables on the overall hazard.

We, therefore, interpret the Cox hazard model using hazard ratios, defined as the ratio of the predicted hazard function of two different values of a predictor variable. When the hazard ratio is less than 1, then the event of interest is less likely to occur, and when the hazard ratio is greater than 1, the event of interest is more likely to occur. When the hazard ratio equals 1, then the predictor variable does not affect the hazard function of the event.

An advantage of the Cox model is that, due to its regression framework, we can estimate the hazard ratios that are controlled for other variables in the model, such as race, sex, and age. Another reason the Cox models have achieved immense popularity is that they don’t require the researcher to assume a specific survival distribution for the data. Instead, the models use a hazard function.

### 3.6.3. Proportional Hazard Property

Despite the seeming ease of the Cox model, there is an important assumption that must be checked before using it. It assumes proportional hazards between predictor values irrespective of how the undying hazard function may change with time.

That is given two different observations \( x_1 \) and \( x_2 \), the hazard ratio is as follows:
\[
\frac{h(t|x_1)}{h(t|x_2)} = \frac{\exp(x_1\beta)}{\exp(x_2\beta)} = \exp(\langle x_1 - x_2 \rangle \beta)
\]

This shows that the hazard ratio of the two observations is constant with respect to time \( t \).

For any Cox proportional hazard model, the survival function for \( T \) is as follows:
\[
S(t|x) = \exp \left( \int_0^t \lambda(u|x) \, du \right)
\]
\[
= \exp \left( -\exp(X'\beta) \int_0^t \lambda_0(u) \, du \right)
\]
\[
= \exp \left( \int_0^t \lambda_0(u) \, du \right)^{\exp(X'\beta)} = (S_0(t))^{\exp(X'\beta)}
\]

where \( S_0(t) \) is the baseline hazard function. Therefore, the probability distribution function of \( T \) is given by:
\[
f(t|x) = \lambda(t|x)S(t|x) = \lambda_0(t) \exp(X'\beta) (S_0(t))^{\exp(X'\beta)}
\]

Note that if the proportionality assumption holds, then
\[
S_0(t) = (S_0(t))^{\exp(X'\beta)}
\]

Since the model is built solely around this assumption, if it turns out to be invalid for several predictors in a given data set, then the model should not be applied to that data set.

A way around this problem is to fit a model that has time-varying covariates or fit a stratified Cox model where the baseline hazard will vary from stratum to stratum.

It’s critical to examine the proportional hazard assumption for every predictor in a Cox model. The best way to do this is to plot the Schoenfeld residuals versus time (this is for continuous variables). For a valid proportional hazard assumption, the Schoenfeld residuals should resemble a random scatter around 0. For categorical variables, it is easier to compare a log-log transformation of the Kaplan Meier survival curves for different categories. For a valid proportional hazard assumption, the survival curves should be almost parallel and should intersect after time apart. There might be a bit of crossing at earlier time points caused by noise in the estimates and may violate the assumption.

### 4. Application to Application to the 2014 Kenya Demographic Health Survey Data

#### 4.1. Introduction to Data

The data used in this study was obtained from the 2014 Kenya Demographic and Health Survey (KDHS). It is a nationwide representative survey of men aged 15-54 and women aged 15-49 selected from 5,360 clusters throughout the country. Both Women’s Questionnaire and Household Questionnaire were administered in all households, while the Men’s Questionnaire was given in every second household.

Just like with any other child mortality data, important information comes from surveys conducted on women. The interviews involved a retrospective maternal history that collected data on every child’s date of birth, survival status, and age at death.

The dependent variable in this study is the duration of time from birth to death or censor. The main explanatory variables are categorized into environmental (drinking water and type of toilet facility), socioeconomic (mothers education, place of residence, place of delivery, and family wealth index), and demographic (family size, sex of a child, mother’s age at birth, preceding interval, breastfeeding habit) factors.

#### 4.2. Descriptive Analysis

From the results, 6,123 (7.3%) of the 83,591 children who were born alive between 2009 and 2014, did not celebrate their fifth birthday. Out of the 83,591 children who were born alive between 2008/9 and 2014, 50.6% were males while 49.4% were females. More than two-thirds (69.8%) of the children were born in rural areas, while only 30.2% were born in urban areas of Kenya. Nairobi region had the least number of children born (1.9%) between 2008/9 and 2014. According to the results, the highest proportion of children (31.3%) were born in the poorest than in the richest households (12.3%). Slightly more than
three-quarters (75.7%) of the children born alive lived with their mothers, 7.6% lived with relatives, and only 6.7% lived with their fathers.

Women who had the lowest level of education (primary or no education at all) gave birth to most children compared to their learned counterparts (secondary and higher education), who gave birth to only 21.7% of the total number of children born alive.

Almost all the children born between 2008/9 and 2014 came from Christian households (81.4%), and only 15.8% came from Muslim households. A higher proportion of children were born from self-employed mothers (83%), with only 3.9% being born of employed mothers.

Over half of the children (64.1%) were born in male-headed households, while only 35.9% came from female-headed households. Out of all children born alive between 2009 and 2014, 6,123 (7.3%) died before the age of five years. A higher proportion of deaths occurred in the urban areas of Kenya.

Most of the deaths (54.6%) occurred among male children (54.6%) than among female children (45.4%). Most of the under-five deaths (70.9%) occurred in rural areas, while only 29.1% of the deaths occurred in the urban areas of Kenya.

A child born in the Nyanza region (136 deaths per 1,000 live births) is twice likely to die before five years as a child born in the central region (58 deaths per 1,000 live births).

The western region recorded the second-highest number of under-five deaths (115 deaths per 1,000 live births).

As expected, the majority (32.6%) of the children who died before attaining the age of five came from the poorest families, with only 8.4% coming from the richest families. Of all the deaths, firstborn children recorded the highest number of deaths (28.2%) compared to other birth orders. Most of the deaths (84.9%) occurred to children whose mothers have attained a lower level of education (primary and below), while only 2.4% occurred to children whose mothers had a higher level of education.

The majority of the children who died before five years came from male-headed households (62.8%) than from female-headed households (37.2%). The results indicate that children born as a result of multiple births recorded the highest number of deaths than those born as a result of singleton births. Of the 2171 multiple births, 450 (20.7%) died before reaching the age of five, while only 6.2% of singleton births died before the age of five. Table 1 presents a summary of the distribution of live births and deaths among under-five children.

| Covariates               | Frequency | Proportion | Child is Dead | Child is Alive |
|--------------------------|-----------|------------|---------------|----------------|
| Sex of Child             |           |            |               |                |
| Male                     | 42,337    | 50.6       | 54.6          | 50.3           |
| Female                   | 41,254    | 49.4       | 45.4          | 49.7           |
| Place or Residence       |           |            |               |                |
| Rural                    | 25,261    | 30.2       | 29.1          | 30.3           |
| Coast                    | 10,350    | 12.4       | 14.1          | 12.2           |
| North Eastern            | 5,738     | 6.9        | 5.9           | 6.9            |
| Eastern                  | 13,109    | 15.7       | 12            | 16             |
| Region                   |           |            |               |                |
| Rift Valley              | 6,678     | 8          | 6             | 8.1            |
| Western                  | 25,367    | 30.3       | 21.8          | 31             |
| Nairobi                  | 8,145     | 9.7        | 13.7          | 9.4            |
| Nyanza                   | 12,635    | 15.1       | 24.7          | 14.4           |
| Nairobi                  | 1,569     | 1.9        | 1.6           | 1.9            |
| No Education             | 18,303    | 21.7       | 21.9          | 21.9           |
| Primary                  | 47,126    | 56.4       | 63            | 55.9           |
| Mother's Educational Level|         |            |               |                |
| Secondary                | 14,220    | 17         | 12.9          | 17.3           |
| Higher                   | 3,942     | 4.7        | 2.4           | 4.9            |
| Poorer                   | 26,170    | 31.3       | 32.6          | 31.2           |
| Poorer                   | 17,926    | 21.4       | 25.2          | 21.1           |
| Middle                   | 15,905    | 19         | 19.4          | 19             |
| Richer                   | 13,316    | 15.9       | 14.3          | 16.1           |
| Richest                  | 10,271    | 12.3       | 8.4           | 12.6           |
| Wealth Index             |           |            |               |                |
| Sex of HH Head           |           |            |               |                |
| Male                     | 53,109    | 64.1       | 62.8          | 64.2           |
| Female                   | 29,794    | 35.9       | 37.2          | 35.8           |
| Religion                 |           |            |               |                |
| Roman Catholic           | 15,859    | 19         | 16.7          | 19.2           |
| Protestants              | 52,214    | 62.5       | 64.4          | 62.3           |
| Muslims                  | 13,163    | 15.7       | 16.1          | 15.7           |
| Other Religion           | 2,355     | 2.8        | 2.8           | 2.8            |
| 15-19                    | 1,080     | 1.3        | 0.5           | 1.4            |
| 20-29                    | 21,210    | 25.4       | 18.3          | 25.9           |
| 30-39                    | 33,026    | 39.5       | 36.6          | 39.7           |
| 40-49                    | 28,275    | 33.8       | 44.7          | 33             |
| Mother's Age             |           |            |               |                |
| Not Employed             | 10,985    | 13.1       | 11.7          | 13.3           |
| Employed                 | 3,227     | 3.9        | 3.7           | 3.9            |
| Mother's Occupation      |           |            |               |                |
| Self Employed            | 69,379    | 83         | 84.7          | 82.9           |
| 1st Birth                | 23,245    | 27.8       | 28.2          | 27.8           |
| 2nd Birth                | 18,819    | 22.5       | 21.7          | 22.6           |
| 3rd Birth                | 14,081    | 16.8       | 15.9          | 16.9           |
| 4th Birth                | 9,963     | 11.9       | 11.7          | 11.9           |
| 5th & Higher Births      | 17,483    | 20.9       | 22.6          | 20.8           |

Table 1. Distribution of births and deaths by survival determinants.
4.3. Non-parametric Survival Analysis: Kaplan-Meier Model

The figures below present the Kaplan-Meier curves for a few selected risk factors for under-five child mortality in Kenya from KDHS 2014 data set.

A male child is at higher risk of dying before attaining the age of five than a female child, as illustrated in the figures below.

From figure 4, Central and Rift Valley regions have lower mortality rates compared to other regions. However, Western and Nyanza regions have high under-five mortality rates. This means that children born in either Western or Nyanza regions experience a higher probability of dying before the age of five than children born in either Central or Rift Valley regions.

It is clear from figure 4 that a child born of a mother with a secondary and higher level of education stands a higher chance of survival than that of a mother with a primary or lower level of education.

The survival curve for children born in urban areas is above the survival curve of children who are born in rural settings. This implies that children born in urban areas are at a higher chance of surviving to the age of five years than children born in rural areas. The wealth of a family affects the survival of a child. A child born from the wealthiest family has a higher probability of survival than a child born from a poorer family. Children who were breastfed were at a lower risk of dying before attaining the age of five years than those who were not.

4.4. Fitting a Standard Cox-proportional Hazard Model on the KDHS 2014 Data

Assuming the time to death of the children under the age of five in the KDHS 2014 data are identically and independently distributed, the Cox Proportional Hazard model with several covariates was fitted. Table 2 presents the unadjusted hazard
ratios, the p-values, and the 95% confidence intervals from fitting the standard univariate Cox-proportional hazard model.

The results indicate that a mother's education is a significant risk factor associated with under-five child mortality. Attending secondary education and above is associated with a reduction in the chances of a child dying before reaching five years. A child born of a mother with a higher level of education had higher chances \((b = -0.063; SE = 0.087; p - value < 0.01)\) of surviving to age five than one born of a mother with secondary \((b = -0.026; SE = 0.045; p - value < 0.01)\) and lower-level \((b = 0.113; SE = 0.032; p - value < 0.01)\) of education. The region where a child is bred and raised affects his/her chances of surviving to the age of five years. The significant positive coefficients for both Western \((b = 0.187, SE = 0.048, p - value < 0.01)\) and Nyanza \((b = 0.352, SE = 0.043, p - value < 0.01)\) regions indicate that the hazard rate for under-five mortality is greater in the two regions than in other regions of the country. This implies that children born in the Western and Nyanza regions are more likely to die before the age of five than those born in other regions. On the other hand, being born in Central, Rift Valley, North Eastern, Eastern, and Nairobi regions was negatively predictive of the hazard for under-five mortality (Central: \(b = -0.428, SE = 0.062, p - value < 0.01\); Rift Valley: \(b = -0.466, SE = 0.044, p - value < 0.01\); North Eastern: \(b = -0.266, SE = 0.062, p - value < 0.01\); and Nairobi: \(b = -0.232, SE = 0.105, p - value = 0.01\)). This means that children born in these regions experience lower mortality rates than those born in the Western and Nyanza part of the country.

There is a strong positive association between place of residence \((b = 0.061, SE = 0.033, p - value = 0.013)\) and the survival of under-five children. The hazard of under-five mortality is greater among children born in rural areas than those born in urban areas. In other words, children born in rural areas of Kenya experience higher chances of dying before attaining the age of five compared to those born in the urban areas of the country.

\[\text{Table 2. The Univariate Cox Proportional Hazard Model.}\]

| Factor                     | B (SE)      | Sig. | Unadjusted HR (95% CI) |
|----------------------------|-------------|------|------------------------|
| Place of Residence         | 0.061 (0.033) | 0.013 | 0.980 (0.970, 0.989) |
| Education                  | 0           |      |                        |
| Education Primary          | 0.113 (0.032) | 0      | 1.119 (0.890, 1.364) |
| Education Secondary        | -0.026 (0.045) | 0    | 0.770 (0.702, 0.972) |
| Education Higher           | -0.063 (0.087) | 0 | 0.532 (0.502, 0.682) |
| Family Size                | 0.017 (0.002) | 0    | 1.017 (1.013, 1.021) |
| Number of Children Under 5 | 0.035 (0.005) | 0 | 1.036 (1.027, 1.045) |
| Total # of Children        | 0.255 (0.008) | 0 | 1.290 (1.269, 1.312) |
| Sex of Child               | 0.007 (0.007) | 0 | 1.007 (0.993, 1.021) |
| Mother’s Age               | 0           |      |                        |
| Mother’s Age (1)           | 1.627 (0.032) | 0 | 5.090 (4.784, 5.516) |
| Mother’s Age (2)           | 0.482 (0.010) | 0 | 1.620 (1.590, 1.650) |
| Mother’s Age (3)           | 0.180 (0.006) | 0 | 1.197 (1.177, 1.217) |
| Delivery Hospital          | 0.232 (0.011) | 0 | 1.262 (1.234, 1.290) |
| Postnatal                  | -0.65 (0.026) | 0 | 1.106 (1.090, 1.123) |
| Breastfeeding              | -0.264 (0.089) | 0.003 | 0.768 (0.645, 0.915) |
| Birth Type                 | 0.063 (0.034) | 0.061 | 1.065 (0.997, 1.138) |
| Preceding                  | -0.014 (0.001) | 0 | 1.006 (1.005, 1.006) |
| Succeeding                 | -0.024 (0.001) | 0 | 1.003 (1.003, 1.003) |
| Religion1                  | 0           |      |                        |
| Religion1 (1)              | 0.165 (0.035) | 0 | 0.942 (0.900, 0.986) |
| Religion1 (2)              | 0.178 (0.045) | 0 | 0.934 (0.894, 0.976) |
| Religion1 (3)              | 0.15 (0.082) | 0.068 | 0.981 (0.935, 1.031) |
| Region1                    | 0           |      |                        |
| Region N. Eastern          | -0.26599 (0.062) | 0 | 0.992 (0.938, 1.050) |
| Region Eastem              | -0.408 (0.050) | 0 | 0.665 (0.612, 0.957) |
| Region Central             | -0.428 (0.062) | 0 | 0.969 (0.918, 1.024) |
| Region Rift Valley         | -0.466 (0.044) | 0 | 0.934 (0.882, 0.990) |
| Region Western             | 0.187 (0.048) | 0 | 1.015 (0.962, 1.071) |
| Region Nyanza              | 0.352 (0.043) | 0 | 0.949 (0.896, 1.005) |
| Region Nairobi             | -0.232 (0.105) | 0.001 | 0.908 (0.859, 0.959) |
| Water                      | 0.120 (0.026) | 0 | 0.951 (0.936, 0.965) |
| Toilet                     | 0.232 (0.027) | 0 | 0.983 (0.968, 0.998) |
| Occupation1                | 0           |      |                        |
| Occupation1 (1)            | 0.066 (0.076) | 0.39 | 1.087 (1.64, 1.110) |
| Occupation1 (2)            | 0.109 (0.04) | 0.006 | 1.047 (1.009, 1.086) |
| BirthOrders                | 0           |      |                        |
| BirthOrders (1)            | -0.019 (0.037) | 0.06 | 0.746 (0.730, 0.761) |
| BirthOrders (2)            | -0.0213 (0.040) | 0.044 | 0.871 (0.853, 0.891) |
| BirthOrders (3)            | 0.02 (0.045) | 0.657 | 0.923 (0.902, 0.945) |
| BirthOrders (4)            | 0.143 (0.036) | 0 | 0.950 (0.926, 0.975) |
| Wealth Index               | 0           |      |                        |
| Wealth Index Poorer        | 0.110 (0.034) | 0.001 | 1.032 (1.001, 1.064) |
The significant positive coefficient \((b = 0.0232, SE = 0.027, p - value < 0.01)\) indicates that the hazard of under-five mortality is greater among children who live in households that use unimproved toilet facilities than among those living in households that use improved toilet facilities. The source of drinking water has a strong association with child mortality. The risk of a child dying before the age of five is higher among children born in households that access unsafe sources \((b = 0.120; SE = 0.026; p - value < 0.01)\) of drinking water than those in a household with access to safe sources of water.

The death of under-five children increases with family size \((b = 0.017, SE = 0.002, p - value < 0.01)\). In other words, children born in families with more family members are more likely to die before attaining the age of five than those born in smaller-sized families. The results show the total number of children under-five in one household was a significant positive \((b = 0.105, SE = 0.005, p - value < 0.01)\) predictor of the hazard for the under-five mortality. This implies that a child born in a family that has a higher number of children who are under the age of five years is more likely to die before attaining the age of five than children born in a family with fewer under-five-year-old children.

The number of total children ever born was a significant negative \((b = 0.255, SE = 0.008, p - value < 0.01)\) predictor of the hazard of under-five mortality. This means that an increase in the number of children ever born in a household increases the chances of a child born in that household dying before attaining the age of five years.

Preceding birth interval was a significant negative \((b = -0.024; SE = 0.001; p - value < 0.01)\) predictor of the hazard of under-five mortality. In other words, increasing preceding birth interval lengths increases the chances of children surviving to the age of five years.

The higher succeeding birth interval was a significant negative \((b = -0.024; SE = 0.001; p - value < 0.01)\) predictor of the hazard of under-five mortality. In other words, increasing succeeding birth interval lengths reduces the chances of a child dying before reaching the age of five children. (multiple versus single birth) is also a significant \((b = 1.291; SE = 0.049; p - value = 0.061)\) predictor of under-five child mortality.

Mother's occupation is strongly associated with the survival of under-five-year-old children. A child born of a mother who is either employed \((b = 0.109; SE = 0.04; p - value = 0.006)\) or self-employed \((b = 0.066; SE = 0.079; p - value = 0.39)\) has a higher chance of dying before the age of five years than one who is born of a mother who is not employed.

The sex of a household head and birth type did not affect the survival of under-five children.

### 4.4.1. Testing Proportional Hazard Assumption Using the Scaled Schoenfeld Residuals

To conduct the multiple covariate Cox proportional hazard model, we need to test for one of the most important assumptions of the model, the proportionality hazard assumption. Not all the covariates presented in Table 2 satisfy the proportionality hazard assumption.

To test for the proportionality hazard assumption, the scaled Schoenfeld residuals test was conducted in R using the cox.zph command. The results of the test are presented in Table 3 below.

| Covariate                  | Chi-square (df) | p-value | Covariate                  | Chi-square (df) | p-value |
|----------------------------|-----------------|---------|----------------------------|-----------------|---------|
| Mother's Age               | 36.3 (1)        | <0.01   | Toilet Facility            | 2.84 (1)        | 0.092   |
| GLOBAL                     | 36.3 (1)        | <0.01   | Religion                  | 15.1 (3)        | 0.14    |
| Family Size                | 1.77 (1)        | 0.18    | GLOBAL                     | 11.8 (3)        | 0.19    |
| GLOBAL                     | 1.77 (1)        | 0.18    | Mother's Occupation       | 3.96 (2)        | 0.14    |
| Total Children in HH       | 2.33 (1)        | 0.13    | GLOBAL                     | 13.9 (2)        | 0.14    |
| GLOBAL                     | 2.33 (1)        | 0.13    | GLOBAL                     | 15.8 (4)        | 0.0033  |
| Number of Children Under 5| 0.573 (1)       | 0.45    | Child is Twin              | 17.9 (1)        | <0.001  |
| GLOBAL                     | 0.573 (1)       | 0.45    | GLOBAL                     | 17.9 (1)        | <0.001  |
| Preceding Birth Interval   | 0.518 (1)       | 0.47    | Sex of Child               | 7.6 (1)         | 0.005   |
| GLOBAL                     | 0.518 (1)       | 0.47    | GLOBAL                     | 7.6 (1)         | 0.005   |
| Succeeding Birth Interval  | 137 (1)         | <0.01   | Wealth                     | 15.8 (4)        | 0.0033  |
| GLOBAL                     | 137 (1)         | <0.01   | GLOBAL                     | 15.8 (4)        | 0.0033  |
| Region                     | 23.6 (7)        | 0.0013  | Place of Delivery          | 35.3 (1)        | <0.001  |
| GLOBAL                     | 23.6 (7)        | 0.0013  | GLOBAL                     | 35.3 (1)        | <0.001  |
| Place of Residence         | 2.96 (1)        | 0.086   | Breastfeeding              | 37.6 (1)        | <0.001  |
| GLOBAL                     | 2.96 (1)        | 0.086   | GLOBAL                     | 37.6 (1)        | <0.001  |
| Education                  | 7.94 (3)        | 0.057   | Postnatal                  | 6.59 (1)        | 0.01    |
| GLOBAL                     | 7.94 (3)        | 0.057   | GLOBAL                     | 6.59 (1)        | 0.01    |
| Source of Water            | 0.88 (1)        | 0.35    | Birth Orders               | 6.47 (4)        | 0.17    |
| GLOBAL                     | 0.88 (1)        | 0.35    | GLOBAL                     | 6.47 (4)        | 0.17    |

From the results of the Schoenfeld test, it is clear that mother's age at childbirth, succeeding birth interval, region,
wealth index, sex of a child, multiple births, place of delivery, breastfeeding habits, and postnatal care do not satisfy the proportionality hazard assumption because their p-values are less than 0.05 significance level from the scaled Schoenfeld residuals test. As such, these factors cannot be befit in the final cox-proportional hazard model to get the adjusted hazard ratios. However, other variables like the family size and the total number of children in a household satisfy this assumption and can, therefore, be fitted in the final cox-proportional hazard model.

4.4.2. The Best Fitting Model Selected by the AIC by Akaike

Using Akaike’s Information Criterion (AIC) by Akaike (1973), the best fitting standard Cox proportional hazard model consisted of eleven risk factors, including family size, the total number of children ever born, number of under-5 children in a household, preceding birth interval, place of residence, mother's level of education, type of toilet facility, religion, mother’s occupation, and birth order. The best fitting model is selected in line with the results of other findings [13].

| Covariates                        | B     | SE    | p-value | HR (95% CI) |
|-----------------------------------|-------|-------|---------|-------------|
| Place of Residence Rural          | 0.115 | 0.030 | 0.000   | 1.122 (1.057, 1.191) |
| Education                         |       |       |         |             |
| Education Primary                 | 0.049 | 0.096 | 0.607   | 1.050 (0.871, 1.267) |
| Education Secondary               | 0.388 | 0.088 | 0.000   | 1.474 (1.239, 1.753) |
| Education Higher                  | 0.209 | 0.092 | 0.023   | 1.232 (1.029, 1.476) |
| Family Size                       | -0.220| 0.007 | 0.000   | 0.803 (0.792, 0.814) |
| Number of Children Under 5        | -0.037| 0.015 | 0.014   | 0.964 (0.936, 0.993) |
| Total Number of Children          | 0.289 | 0.006 | 0.000   | 1.335 (1.319, 1.352) |
| Preceding                         | -0.009| 0.001 | 0.000   | 0.991 (0.989, 0.993) |
| Toilet Unimproved                 | -0.143| 0.029 | 0.000   | 0.867 (0.819, 0.917) |
| Religion                          |       |       |         |             |
| Religion1 Protestant              | -0.033| 0.083 | 0.688   | 0.967 (0.821, 1.139) |
| Religion1 Muslim                  | 0.124 | 0.080 | 0.118   | 1.133 (0.969, 1.324) |
| No Religion                       | 0.271 | 0.083 | 0.001   | 1.311 (1.113, 1.543) |
| Occupation                        |       |       |         |             |
| Occupation Employed               | -0.060| 0.041 | 0.146   | 0.942 (0.868, 1.021) |
| Occupation Self-Employed          | 0.189 | 0.071 | 0.008   | 1.208 (1.052, 1.387) |
| Birth Order                       |       |       |         |             |
| Birth Orders 2nd                  | 0.603 | 0.041 | 0.000   | 1.827 (1.687, 1.980) |
| Birth Order 3rd                   | 0.442 | 0.042 | 0.000   | 1.555 (1.433, 1.688) |
| Birth Order 4th                   | 0.326 | 0.044 | 0.000   | 1.386 (1.272, 1.510) |
| Birth Order 5th and Higher        | 0.236 | 0.047 | 0.000   | 1.267 (1.155, 1.389) |

5. Discussion

The goal of this study was to determine the risk factors of under-5 mortality in Kenya. From the results, 6,123 (7.3%) of the 83,591 children born alive between 2009 and 2014 did not celebrate their fifth birthday. The results indicate that family size, the total number of children ever born, number of under-5 children in a household, preceding birth interval, place of residence, mother’s level of education, type of toilet facility, religion, mother’s occupation, and birth order were the determinants of U5CM in Kenya. These results do not deviate much from the findings of studies in other Sub-Saharan African countries such as Ethiopia, Uganda, Tanzania, and Zimbabwe [13].

A child who is born and raised in an urban setting was less likely to die than one who is born and raised in a rural setting. This result is in agreement with the results of other findings [13]. In Kenya, inconsistent distribution of socio-economic resources between rural and urban areas and political influence directly affect health conditions [20]. These findings might also be attributed to the better quality of the health environment and sanitation in urban areas compared to the rural areas. Mothers who had more children in the last five years were more likely to die than one who is born and raised in a rural setting. This result is in agreement with the results of other findings [13].

In Kenya, inconsistent distribution of socio-economic resources between rural and urban areas and political influence directly affect health conditions [20]. These findings might also be attributed to the better quality of the health environment and sanitation in urban areas compared to the rural areas. Mothers who had more children in the last five years were more likely to die than one who is born and raised in a rural setting. This result is in agreement with the results of other findings [13].

Children who live in households that access improved toilet facilities experience a lower risk of death. This result is in line with the findings of the Centers for Disease Control and Prevention, which indicate that approximately 801,000 children below the age of 5 years die from diarrhea each year, mostly in developing countries, as poor hygiene conditions.

Discussion

The results also indicate that short birth intervals are associated with maternal depletion syndrome as well as resource competition between siblings, in addition to a lack of care and attention between rural and urban areas and political influence directly affect health conditions [20]. These findings might also be attributed to the better quality of the health environment and sanitation in urban areas compared to the rural areas. Mothers who had more children in the last five years were more likely to die than one who is born and raised in a rural setting. This result is in agreement with the results of other findings [13].

In Kenya, inconsistent distribution of socio-economic resources between rural and urban areas and political influence directly affect health conditions [20]. These findings might also be attributed to the better quality of the health environment and sanitation in urban areas compared to the rural areas. Mothers who had more children in the last five years were more likely to die than one who is born and raised in a rural setting. This result is in agreement with the results of other findings [13].

Children who live in households that access improved toilet facilities experience a lower risk of death. This result is in line with the findings of the Centers for Disease Control and Prevention, which indicate that approximately 801,000 children below the age of 5 years die from diarrhea each year, mostly in developing countries, as poor hygiene conditions.
Relative to those born in the Central region, children born in and Western and Nyanza regions were more likely to die before the age of five years. This finding may also be attributed to the inconsistent distribution of socio-economic resources across the regions. Besides, due to the proximity to the capital city, Nairobi, which has better health care systems, mothers in the Central region can easily access quality health care services leading to their children experiencing less likelihood of dying before the age of five than those in Nyanza and Western regions.

An unexpected finding was that the risk of under-5 child mortality was not affected by the sex of a child. This is inconsistent with the findings of other similar studies [21]. U5CM is reduced if mothers are more educated (high school and above). This result is in line with other findings [22]. A highly educated mother is more knowledgeable and able to make well-informed decisions about accessing healthcare services, including maternal and child healthcare-seeking behaviors [23], such as vaccinations, which could increase the survival chances of a child. Also, a well-informed mother is likely to reside in an economically developed area with an improved healthcare system and better sanitary conditions, which will enhance the survival probability of her children [24].

6. Conclusion and Recommendation

The study suggests that under-five child mortality is still a problem in Kenya, and the findings should inform decision-makers in resource allocation towards reproductive, maternal, new born, and child health care. The under-5 mortality rate in Kenya was found to be strongly associated with the mother’s level of education, region, place of residence, birth order, family size, mother's occupation, the total number of children, and type of toilet facility. Intensified and effective interventions are required to address these factors to achieve the Sustainable Development Goals by 2030.

Targeted interventions on women empowerment through education and health education, especially in rural and lesser inaccessible areas, will improve the socioeconomic status, health awareness, and health service utilization and ultimately reduce U5CM rates. The national government should focus on implementing socio-economic development interventions that focus on reducing regional disparities. Interventions should focus on promoting the utilization of modern family planning methods for spacing to help reduce the U5CM in Kenya.

Additionally, there is a need for intensified efforts to promote sanitation and hygiene, particularly in rural areas and lesser accessible regions. Strategies aimed towards ensuring equitable regional socioeconomic development may further impact child survival. More studies on health-related factors that impact child mortality are required. Such studies can inform approaches and interventions that are most relevant to the survival of children in Kenya.

Acknowledgements

The completion of this research project has been possible through the help of many individuals who gave their support. I thank the Almighty God for granting me the knowledge to carry out this study. I can’t overlook the love and support from my family and friends who encouraged me throughout the project. Special thanks to my supervisor Dr. Anthony Wanjoya for the scholarly and professional assistance and his presence to correct my work. I highly appreciate the moral and academic support he gave me.

References

[1] Bay G, Miller T, Faijer DJ. WHO: Levels and trends in child mortality. Report 2013. 2013; 30.
[2] Hill K, Bicego G. Childhood Mortality in Kenya: An Examination of Trends and Determinants in the Late 1980s to Mid 1990s.
[3] Hamel MJ, Adazu K, Obor D, et al. A reversal in reductions of child mortality in Western Kenya, 2003-2009. Am J Trop Med Hyg [Internet]. The American Society of Tropical Medicine and Hygiene; 2011 [cited 2021 Jun 10]; 85 (4): 597–605. Available from: /pmc/articles/PMC3183762/.
[4] Schultz TP. Studying the impact of household economic and community variables on child mortality. Child Surviv Strateg Res. Cambridge University Press; Population & Development Review, supplement to v. 10; 1984: 215–235.
[5] Mosley WH, Chen LC. An analytical framework for the study of child survival in developing countries. Child Surviv Strateg Res. Cambridge University Press; Population & Development Review, supplement to vol. 10; 1984: 25–45.
[6] Rutstein SO. Factors associated with trends in infant and child mortality in developing countries during the 1990s [Internet]. Bull. World Health Organ. 2000. Available from: http://www.who.int/bulletin/tableofcontents/.
[7] Davanzo J, Razzaque A, Rahman M, et al. The Effects of Birth Spacing on Infant and Child Mortality, Pregnancy Outcomes, and Maternal Morbidity and Mortality in Matlab, Bangladesh. THE EFFECTS OF BIRTH SPACING ON INFANT AND CHILD MORTALITY, PREGNANCY OUTCOMES, AND MATERNAL MORBIDITY AND MORTALITY IN MATLAB, BANGLADESH *. 2004.
[8] Kibet MK. Comparative study of infant and child mortality: The case of Kenya and South Africa. African Popul Stud [Internet]. Stellenbosch University; 2013 [cited 2021 Jun 10]; 24 (1–2): 1–2. Available from: http://aps.journals.ac.za.
[9] Rutstein SO. Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: Evidence from the demographic and health surveys. Int J Gynecol Obstet [Internet]. John Wiley and Sons Ltd; 2005 [cited 2021 Jun 10]; 89 (SUPPL. 1): S7–S24. Available from:https://obgyn.onlinelibrary.wiley.com/doi/full/10.1016/j.ijigo.2004.11.012.
[10] Kumar P, File G. Infant and child mortality in Ethiopia: A statistical analysis approach. Ethiop J Educ Sci [Internet], African Journals Online (AJOL); 2011 [cited 2021 Jun 10]; 5 (2). Available from: https://www.ajol.info/index.php/ejesc/article/view/65373.
[11] Mutunga CJ. Environmental Determinants of Child Mortality in Kenya. Heal Inequal Dev [Internet]. Palgrave Macmillan UK; 2011 [cited 2021 Jun 10]. p. 89–110. Available from: https://link.springer.com/chapter/10.1057/9780230304673_5.

[12] Wardlaw T, Jensen L, Moreira Daniels webSite L, Cassidy S, Yuen D, Rao Gupta G. Committing to Child Survival: A Promise Renewed UNICEF Country Offices contributed to the review of country example text. Policy and communications advice and support were provided by [Internet]. 2012 [cited 2021 Jun 10]. Available from: www.apromiserenewed.org.

[13] Hobcraft JN, Mcdonald JW, Rutstein SO. Socio-economic factors in Infant and child mortality: A cross-national comparison. Popul Stud (NY). 1984; 38 (2): 193–223.

[14] Stallings RY. Child morbidity and treatment patterns [Internet]. 2004 Dec. Available from: http://www.measuredhs.com.

[15] Ngome E. Demographic, Socio-Economic and Psychosocial Determinants of Current and Consistent Condom Use among Adolescents in Botswana. World J AIDS. Scientific Research Publishing, Inc; 2016; 06 (04): 137–156.

[16] Mutunga CJ. A Service of zbw Environmental determinants of child mortality in Kenya Standard-Nutzungsbedingungen [Internet]. [cited 2021 Jun 10]. Available from: www.econstor.eu.

[17] Uaps 2007 [Internet]. [cited 2021 Jun 10]. Available from: https://uaps2007.princeton.edu/abstracts/70670.

[18] Fayehun OA. Household Environmental Health Hazards and Child Survival in Sub-Saharan Africa [WP74] [Internet]. 2010. Available from: www.measuredhs.com.

[19] Buttenheim AM. The sanitation environment in urban slums: Implications for child health. Popul Environ. 2008; 30 (1–2): 26–47.

[20] Ilincs S, Giorgio L Di, Salari P, Chuma J. Socio-economic inequality and inequity in use of health care services in Kenya: Evidence from the fourth Kenya household health expenditure and utilization survey [Internet]. Int. J. Equity Health. BioMed Central Ltd.; 2019 [cited 2021 Jun 10]. p. 1–13. Available from: https://doi.org/10.1186/s12939-019-1106-z.

[21] Malderen C Van, Oyen H Van, Speybroeck N. Contributing determinants of overall and wealth-related inequality in under-5 mortality in 13 African countries. J Epidemiol Community Health. 2013; 67 (8): 667–676.

[22] Nasejje JB, Mwambi HG, Aphia 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 Biostatistics and methods. BMC Public Health [Internet]. BioMed Central Ltd.; 2015 [cited 2021 Jun 10]; 15 (1): 1–12. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-015-2332-y.

[23] Basu AM, Stephenson R. Low levels of maternal education and the proximate determinants of childhood mortality: A little learning is not a dangerous thing. Soc Sci Med [Internet]. Elsevier Ltd; 2005 [cited 2021 Jun 10]; 60 (9): 2011–2023. Available from: https://pubmed.ncbi.nlm.nih.gov/15743650/.

[24] Ezeh OK, Agho KE, Dibley MJ, Hall J, Page AN. Determinants of neonatal mortality in Nigeria: Evidence from the 2008 demographic and health survey. BMC Public Health [Internet]. BioMed Central Ltd.; 2014 [cited 2021 Jun 10]; 14 (1): 1–10. Available from: http://www.biomedcentral.com/1471-2458/14/521.

Biography

Elda Naliaka Watulo is a statistician with expertise in statistical data analysis, database development and management, and report writing. She has worked with AMREF Health Africa, Kenya and Deutsche Stiftung Weltbevölkerung (DSW) as a data, monitoring and evaluation specialist. She has practical experience in designing and conducting impact evaluations, conducting both qualitative and quantitative data analysis using SPSS, STATA, R, Advanced MExcel, and Nvivo. She has published several papers in the Journal of Water, Sanitation & Hygiene for Development, and others.