ESTIMATION OF CHILDHOOD MORTALITY IN KENYA USING THE 2015/2016 KENYA INTEGRATED HOUSEHOLD AND BUDGET SURVEY

KONGO SCHOLASTICA KINGI

Q56/6912/2017

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE DEGREE IN POPULATION STUDIES, UNIVERSITY OF NAIROBI

DECLARATION

This is to declare that this is my original work and has not been presented to any other university for my academic award.

Signature ....DocuSigned by: Scholastica Kingi E17E290CA568485... Date ...11/19/2020...

Kongo Scholastica Kingi

Q56/6912/2017

This project has been submitted with my approval as the university supervisor.

Signature ....DocuSigned by: Otieno Alfred F2A7FC61698540D... Date ...11/19/2020...

Prof. Alfred O. Agwanda

Population Studies and Research Institute

University of Nairobi

Signature ....DocuSigned by: Anne Khasakhala 4E06B9FE43FB452... Date ...11/19/2020...

Dr. Anne Khasakhala

Population Studies and Research Institute

University of Nairobi

DEDICATION

This research project is dedicated to my loving family, friends, colleagues and to all those who have inspired and encouraged me in diverse ways to strive to achieve my objective. Their encouragement and patience in my absence while putting together the project gave me more strength and willpower to bring forth this document.

ACKNOWLEDGEMENT

I acknowledge the Almighty God for good health, infinite providence and the benevolent interventions even when it seemed difficult and beyond my mortal capabilities.

My heartfelt appreciations go to my supervisors Prof. Alfred O. Agwanda and Dr. Anne Khasakhala for understanding, patience and special guidance and the many hours spent in supervising me through the journey of putting together this work. Their valuable inputs cannot go unnoticed.

Second, all my lecturers through whose hands, I have been changed to view development from the population and social perception. My gratitude goes to Mr. Ben Jarabi, Prof. Kimani Murungaru, Dr. George Odipo, Dr Wanjiru Gichuhi, Prof Lawrence Ikamari, Dr. Samwel Wakibi, Dr. Andrew Mutuku and Dr. Boniface Koyugi.

My special thanks go to my family for moral and financial support without which this journey would have been impossible.

Appreciation also goes to friends, classmates and colleagues for being supportive and inspirational in the course of preparation of this proposal.

Recognition also goes to the UoN and KNBS fraternity for all facilitation towards achievement of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGUREs

ABSTRACT

Childhood mortality is one of the key indicators used in determining both the general and national socio-economic development, such as resource allocation and policymaking, among other uses. This can only be achieved by ensuring that data are accurate at both the national and county level. This study aimed to estimate childhood mortality in Kenya using the 2015/2016 Kenya integrated household and budget survey (KIHBS). The study specifically sought to establish the quality of KIHBS data, levels of under-five mortalities at the county level, and whether the KIHBS estimates on childhood mortality at county levels are plausible. A survey research design was adopted in this study. Secondary data from the 2015/2016 Kenya Integrated Household Budget Survey (KIHBS) and statistics on child deaths and childbirths between September 2015 and August 2016 was used. The survey design deliberately captured a family range of demographic data such as age, child deaths and childbirths, and women with the highest educational level. Data quality was checked using the age ratio method, Whipples, and United Nations accuracy score method. The differential proportion of children dead by the women's age was also assessed to establish the pattern across age groups. Coale–Demeny West model was used to calculate the under-five mortality. Regression was used to check the plausibility of U5MR from the 2015/2016 Kenya Integrated Household Budget Survey (KIHBS). The study found that the 2015/2016 Kenya Integrated Household Budget Survey (KIHBS) data was inaccurately reported with large fluctuations in age ratios for males and females, indicating persons of various ages being carried across age group boundaries or persons misreporting their ages for various reasons. The study concludes that quality issues exist in the 2015/16 KIHBS data. It is not suitable to estimate the under-five mortality rate at the county level because it did not have recent fertility (births in the last 12 months) and had inconsistent results for all the counties. The study recommends that agencies involved in data collection should engage the public in understanding the importance of accurate reporting by creating policies aligned to supporting mass awareness and educating people on how false reporting during surveys distorts information derived from data and how their lives are likely to be affected.

CHAPTER ONE: INTRODUCTION

1.1 Background Information

Childhood mortality is one of the key indicators used in determining both the general and national socio-development. In the past ten years, the international community has set and periodically reviewed child mortality targets (Ferrarini and Norström, 2010). Under-five mortality has regularly been renewed as part of the Sustainable Development Goals (SDGs) pioneered by the United Nations (United Nations Department of Public Information, 2015). Among the 17 SDGs, the third one is concerned with the reduction of neonatal mortality. The objective of this goal is to reduce neonatal deaths to 12 per 1000 live births. It also aims to reduce under-5 mortality to less than 25 per 1000 live births (Pillay, 2019). Mortality is a significant indicator of a society's wellbeing. There have been many social and economic developments in the 21st century that impact social services. Lately, there has been a general decrease in mortality rates in the world partially due to improvements in standards of living, sanitation, advancements in employment, provision of healthcare and sanitation facilities, availability of safe water, as well as the provision of affordable housing (Romani & Anderson, 2002).

Child death is a considerable burden in Africa, where most people are young (WHO, 2005; World Bank, 2006). Childhood mortality varies across the globe, but it is highest in sub-Saharan Africa and South Asia (Heckmann, 2015). For instance, more than 50 percent of childhood deaths occurred in sub-Saharan Africa in 2017. Further, Southern Asia accounted for 30 percent of global childhood mortality in the same year. United Nations Inter-agency Group for Child Mortality Estimation (UNIGME) (2018) found 38 percent of all childhood deaths occur in countries listed among the least developed. High childhood deaths in sub-Saharan Africa may be owed to an increase in fertility in this region. Childhood deaths in the area have increased from 30 percent in 1990 to 50 percent in 2017. Projections indicate that sub-Saharan Africa will be accounting for 60 percent of global childhood deaths by the year 2050 (Renschler et al., 2015).

Several World Health Organization publications have focused on the issue of childhood mortality. This includes a WHO report in 2005 which highlighted the causes of childhood mortality. The report linked childhood deaths to several conditions (World Health Organization, 2005). While progress is being made in fighting child mortality, far too many children do not live to the age of five. UNIGME (2018) reported around 5.4 million deaths among children below the age of 5 in 2017 alone. Almost 50% of the deaths took place in Africa.

1.2 Problem Statement

Estimation of U5MR is highly dependent on data from vital registration systems containing national reported birth and death histories. However, most developing countries have little or no records of such data, yet many childhood deaths occur in these nations. Data to estimate U5MR is instead provided for by full or summary birth histories. The 2014 Kenya Demographic and Health Survey is the latest data source on complete and summary birth histories.

Estimation of under-five mortality is essential. The estimates so obtained are specific to the population and are used for programmatic policy, planning, and research for socio-economic development. Consequently, owing to its importance, KIHBS data should be reported as precisely or concisely as possible. Otherwise, any misreporting can adversely affect various demographic measures. For example, reported age should be accurate to ensure estimates such as neonatal and child mortality are plausible. In contrast, females' ages must be as precise as possible to avoid unnecessarily transferring women into reproductive years, which may, in turn, affect various fertility estimates. Similarly, programmatic planning and its budgeting for children will require flawless data on age to inform on who in the population falls into these categories.

However, data on U5MR in Kenya has had quality issues. Age is often misreported because many people do not know their ages precisely, which implies that the reported age-sex distributions are likely to be distorted. This is manifested in age heaping around digits 0 and 5. Among the factors contributing to age heaping is; overstatement and exaggeration of ages of young children and males and females. While analytical quality checks on the accuracy of reported ages are done for each census, no such efforts are evident for the 2015/2016 KIHBS data, an exercise that is necessary to authenticate estimates so obtained.

Childhood mortality rates in Kenya differ greatly depending on the region where a child is born. The fewest deaths in the under-five-year demographic occur in the central province, where there are 42 deaths in 1000 live births. Conversely, Nyanza recorded the highest number of under-five years of mortality, with as many as 82 deaths per 1000 live births (Mwangi and Murithi, 2015). Nairobi region also has high childhood mortality rates as it ranks second among the regions with high childhood mortality within Kenya. This finding is based on the latest available data. The plausibility of childhood mortality in the counties is also determined. This study sought to establish whether data obtained through the KIHBS 2915/2016 is relevant enough to give plausible estimations of the U5M mortality at the county level in Kenya. This study aims at estimating the under-five mortality rate in Kenyan counties in which these statistics will come from. Therefore, a comparison of the U5M across different counties is made possible.

1.3 Research Questions

These research questions will be vital in addressing the issues mentioned above:

a) What is the quality of KIHBS data?
b) What are the levels of under-five mortality at the county level in Kenya?
c) Are the estimates of U5M at county level plausible from KIHBS data?

1.4 Objectives of the Study

This study will focus on the quality of data on childhood mortality rates in Kenya at the counties.

The specific objectives are:

1. To determine the quality of KIHBS data.
2. To estimate U5M at National and county levels using partial-birth history based on the KIHBS data.
3. To examine the extent to which data from KIHBS is plausible for estimating U5M at the county level.

1.5 Justification

Information on U5MR reporting is pivotal to effective development planning, programming and research, as well as in policy formulation. Indeed, planning by and for private and public sectors, such as community organizations, and services like health and education programmes require separate age by sex data. Information on mortality can be used for research and planning purposes. Mortality statistics are utilized for health planning, population projection, assessment of the failure or success of specific service provision, among other uses. Mortality statistics are also forerunners for epidemiological research on cause of mortality. Policymakers believe that fertility can be reduced by reducing childhood mortality. Therefore, an evaluation of the extent of distortions in reported U5MR data will inform users of the data of its limitations and guide future censuses and surveys and specifically in our case, future KIHBS in Kenya. Further, tabulations on U5MR required in the computation of simple measures relating to population change factors, in the analysis of economic dependence studies. However, childhood mortality rates are significantly higher in third world countries and are caused by factors that differ from those that cause death later in life. Measurements of childhood mortality are vital indicators of the socioeconomic status of a nation or region.

In directing global measures to enhance child survival, up-to-date data and knowledge on patterns and causes of child deaths are critical. Ending preventable deaths of babies and children under five years of age is a priority under the Sustainable Development Goals (SDGs). Under the developmental objectives, reducing infant mortality to as low as 25 cases per 1000 live births is projected to be accomplished when quality data is available to understand patterns and focus areas. Similarly, to obtain reliable time series estimates of under-five mortality rates depends on the data collected in both censuses and surveys.

The complete recording of deaths in the death registration systems in all countries is vital to achieving the SDG child mortality reduction goals. However, the high instances of child mortality in the African countries are a testament to the lack of quality prevention programs that should arise from forecasts and estimates produced by updated and quality data. This study is essential in realizing the gap by using available data sources and determining whether they are plausible in the estimation.

1.6 Scope and Limitations

This study focuses on the estimation of the under-five mortality at the county level in Kenya. The study uses data from 2015/2016 KIHBS to determine the mortality rate of children 5 years or below. In so doing, this study identifies whether the available sources provide data that can give plausible estimations.

1. Specific birth rates were not available; hence were borrowed from published data of the KDHS 2014.
2. The study worked on the assumption that estimates would not change much, yet the time difference between surveys was significant.

CHAPTER TWO: LITERATURE REVIEW

2.1. Introduction

This section gives an overview of the previous literary works on child mortality. The chapter provides the framework on the topic for the specific case study comprising of the principal research focus. The context of the review of previous research studies is set by providing: the purpose of the review for the research's case study, remarks on different methods which have used in the estimation of the under-five mortality in the world and also specific to developing countries, and how the scope of the work is presented in this chapter.

The trends and determinants of childhood mortality are discussed in this chapter with a broader perspective of African nations. Hodge and Jimenez-Soto (2013) suggest that the globe has seen a significant decline in child mortality. Sub-Saharan countries have seen a 50 percent decline in child mortality since 1990. However, UNIGME (2018) reports that majority of child deaths occur in sub-Saharan Africa. Similarly, projections of the UN indicate that more than half of the world's under-five deaths will be happening in sub-Saharan Africa by 2050 (Okiro and Ayieko, 2018). Research studies need to be undertaken to determine the trends, causes, and mitigation strategies against child mortality in less-developed countries (Okiro and Ayieko, 2018). The procedures for the direct and indirect approaches in estimation under-five mortality rates are discussed in this chapter.

2.2 Estimation of Childhood Mortality

Worldwide, the well-being of any nation's children depends on under-five mortality rate (U5MR) as the significant barometer. Specifically, it indicates the socioeconomic progress of any country. Generally, U5MR represents the number of deaths per thousand live births. This epitomizes the risk that an infant born in a particular year will die before reaching the age of five if existing age-specific mortality rates are affected (UNIGME, 2018). Millennium Development Goal 4 (MDG 4) of the United Nations focuses on assessing the success of the nation and monitoring the development of infant mortality reduction by using U5MRR national estimates. According to a survey conducted in 1990-2015, there was a significant two-thirds reduction in U5MR globally (UNIGME 2018), equivalent to reduction of 4.4% per annum. It becomes challenging to estimate the trends and levels in U5MR due to the absence of well-functioning vital registration systems in the great majority of developing countries. This is not only due to lack of available data but also poor data quality. The United Nations Inter-agency Group for Child Mortality Estimation (UNIGME, involving United Nations Population Division, World Bank, World Health Organization and United Nations Children's Fund) publishes as well as generates the valuations of child mortality comparable around the nations including the every year data of 194 countries.

2.2.1 Data Sources

Mortality rates can be estimated in various ways depending on the source of data. Most of the organizations apply a standard method in evaluating under-five mortality across countries to obtain valid comparisons. Data are collected from censuses, surveys and administrative sources such as vital registration. When using essential registration, mortality rates are derived from an abridged life table having a standard period that uses age-specific deaths and mid-year populations. Census and survey data on mortality of children take two different ways: The complete history of birth in which women are asked the date of birth of each child they have born, if the children are alive, and the age of death otherwise and the summary history of birth in which women are asked about the number of children they have ever born and those who have died. Either method results in many projections of infant mortality that apply to some period before the date of the survey. Life tables are used to calculate childhood mortality. There are two main categories of lifetables: abridged (period) life tables and model life tables. Model life tables were developed on empirical studies of age-specific death patterns. They are commonly used to calculate demographic parameters for nations having limited data. life tables give the relative probabilities at specific ages. A model life table also offers different levels of mortality corresponding with life expectancies at birth. The desired characteristic, such as under-five mortality of the corresponding life table, is found by observing the table at the particular level in the model life table. The tables use stable population data to illustrate relationships among different variables. Stable populations give the ultimate effects of a fertility and mortality schedule.

The application of a model life tables assumes that the age-mortality pattern of the studied population is similar to that of the life tables. Besides, population projections employing the factor $5L_x + 5/L_x$ assumes that the age distributed within every five-year interval of the population that is stationary is the same as that of the projected population. One major limitation of using model life tables is that people in a population are from different cohorts containing varying mortality experiences. In contrast, information from various cohorts is as if combined into one table. The disparities in the trends mortality across cohorts can affect life table values and, therefore, provide excess mortality measures.

2.3 Techniques to Estimate Child Mortality

According to literature, the two main approaches used in estimating child mortality are the direct and indirect methods. Data on the date of birth of children, the state of their survival, the dates when deaths occur, or the ages of deceased children are used through direct techniques (Kovsted et al., 2002). Indirect strategies, on the other hand, use information on the survival status of children in unique cohorts of mothers. The cohorts could be identified typically with age or the time since the mothers' first birth. Direct methods use data obtained from specifically designed surveys containing histories of childbirth and vital statistics (Uddin et al., 2009). Indirect methods can use data collected from censuses and other general surveys (Uddin et al., 2009). This study cites the procedures to estimate child mortality documented in DHS publications (Sullivan et al., 1994; Rutstein and Rojas, 2006).

2.3.1 Direct Estimation

Direct approaches to the assessment of infant mortality based on household surveys use birth or pregnancy histories to collect data to assess indicators. Information on each birth or pregnancy of the respondent includes:

- Year and month of birth for each child;
- The sex of every child;
- Whether the child is alive or dead;
- The age of each child who is alive;
- Date or age at death of each dead child;
- The outcome of each pregnancy, such as stillbirth, miscarriage, live birth.

In complete or truncated forms, the compilation of birth or pregnancy histories is completed. Both live births and deliveries by the woman interviewed by the survey date are included in the detailed stories. The data is collected chronologically from the first pregnancy to the last.

Truncated histories include pregnancies or births that have occurred during a certain fixed period. The data is collected reverse chronologically from the date of the survey. The birth history information is arranged in data files having each record to describe a single birth. Other vital variables are the sample weight and the date of the interview.

***Direct Estimation Variants*;** The three essential variants of the direct approaches for computing childhood mortality include:

1. A method based on vital statistics that divides the number of deaths of children under 12 months of age within a particular time span by the total number of births over the same period.
2. A method that includes a reliable cohort life table by dividing the deaths of children under 12 months from a given cohort of births by the total births in that cohort.
3. An approach based on a synthetic cohort life table where mortality probabilities are aggregated into standard age segments for small age segments based on actual cohort mortality.

Calculation Algorithm: The calculation algorithm used in the direct methods tabulates the numerator first, then followed by the denominator. Death is tabulated by identifying the lower limit of the age group (a_1) that the child belonged at the time of death ($a_1 \leq a_i < a_u, where\ a_i\ is\ the\ child's\ age\ at\ death$) and ($t_x - t_y$) is the period in which the child occurred. In particular, ($t_x \leq t_j < t_y$) $where\ t_j\ is\ the\ month\ in\ which\ the\ child\ reached\ the\ lower\ bound\ of\ the\ age\ group\ a_i$.

The algorithm will then tabulate the denominator representing the exposure that the child has witnessed in his/her life. This applies to the exact child from the time of birth to death of the child (if the child died before 5 years of age), then to 5 years of age (if the child is 5 years of age or older or death occurred at 5 years of age or older) or to the date of the interview (if the child is less than 5 years of age). The algorithm specifies the duration reached by each age group in which the child was alive. It is done to determine whether the upper limit of the age group is in the same time or the next. After tabulating the numerators and denominators based on the age group and period, calculation of component probabilities of death in each age group (P_x) for each time frame is done by the division of numerators by the denominators. The under-five mortality rate ($5q_0$) is calculated as follows:

$$U5MR = 1 - \prod_{x}(1 - q_x)$$

Where x represents the age groups 0, 1-2, 3-5, 6-11, 12-23, 24-35, 36-47, 48-59.

2.3.2 Indirect Estimation

Indirect estimation techniques use secondary data from the recorded number of children born and those surviving or dying. There are two types of indirect methods: model-based technique obtained from Brass (Brass & Coale, 1968), and the empirical approach developed by Rajaratnam et al. (2010) at IHME. The Brass method utilizes the mother's age in approximating the child's average length of exposure to a risk of dying. The three variants of the indirect methods are:

1. Estimation of child mortality categorized by the mother’s age (AGE).
2. Estimating child mortality categorized by the duration of marriage (DOM).
3. Estimating child mortality categorized by the time since the first birth (TSFB).

The data needed for the indirect estimation are:

I. The number of women divided into classes of five years of age

II. Marriage time

III. The number of children alive in the 5-year age group

IV. The number of children born alive who died during the five-year age group before or at the time of the survey

V. Number of births within the five-year age group before the survey

In addition, each survey required some of the significant assumptions. Developing the techniques according to trending model can sufficiently denote such assumptions based on under five mortality rate as well as population age trends of fertility. There is no correlation exists in between some factors such as relationship between survival of mother in the population and mortality risk of children, children death does not vary by five-year group of mothers and any time frame. Thus, such factors are enough to indicate the requirements of proper definitions of cohort outlines for childbirth.

1. Ever-Born Girls
   i) The total number of children ever born or born
   ii) The total number of children ever born.
2. Children who survive
   i. The entire number of living or surviving children;
   ii. The total number of surviving children
   iii. The share of surviving children
   iv. Number of dead children, or number of dead children
   v. The average number of children killed or killed, or
   vi. Percentage of dead children
3. Number of females
   i. The total number of women, including women who have never been married (AGE)
   ii. The number of women ever married (DOM)
   iii. The number of women giving birth (TSFB).

Calculation algorithm

The U5MR estimation uses the Loess regression model in 2012 (UNIGME, 2018). The default setting for every nation based on the smoothness parameter that is calculated by data type/availability in the country. The uncertainty assessment of U5MR estimation achieved by using bootstrap approach (Alkema & New, 2012). However, it has many limitations. First issue was with the countries subset where it is necessary to change the value with post-hoc adjustments and the fitted Loess curve was believed to not fit the data appropriately. Second problem was the equal weight of all observations to attain the data quality indicators, potential data biases, standard errors and point estimates that were not responsible for. Calibrating the uncertainty intervals as well as resulting point estimates can enhance the quality of findings.

Therefore, the Institute for Health Metrics and Evaluation (IHME) developed the alternative approaches to evaluate the child mortality for all nations, e.g. Gaussian process regression model. This process also has some room for improvements as the model validates the performance based on 2010 IHME version, possibly not completely responsible for potential data biases. This chapter highlights the current indirect U5MR estimation approach to enhance upon the limitations occurred in previous methodologies due to insufficient calibration in existing methods. One of the current model for data assessment is B3 model including Bias-reduction B-spline and Bayesian approaches. By far, UN IGME has decided to adapt the B3 model for analyzing the development of nations towards the evaluations of B3 and MDG 4 involving in the “Child Mortality Report 2018” (UNIGME, 2018).

The steps in computing the indirect estimates are as follows:

1. Sum the number of children ever to be born and those surviving by age group, or marriage duration, or the period since first birth.

2. Compute the mean parity per woman by group:

$P_i = CEB_i / N_i$ Where P_i is the parity for age group i, CEB_i represents the number of children who have ever been born for age group i and N_i represents the total number of women that particular age group.

3. The proportion of dead children by the group is calculated:

$D_i = (CEB_i - CS_i / N_i)$

4. Multipliers are calculated then followed by the probabilities of dying, which are computed as shown below:

$q_x = k_i * D_i$ Whereby q_x represents the probability of dying at age x exactly.

5. The reference period is calculated, and finally, the values of q_x for each group are transformed into $1q_0$, $4q_1$, and $5q_0$ for every reference point. The transformation is made possible by

applying linear interpolation between the levels of the model life table. TheU5MR is then calculated by:

$$q_5 = q_{5+1} + \theta * (q_{5i} - q_{5+1})$$

Assumptions of the Indirect Estimation;

- The respondent correctly records data about the CEB and CS.
- Awareness of fertility and modes of mortality exists.
- In the last 15 years, fertility and mortality levels have been stable.
- Conditions of mortality are homogeneous - that is, similar risks of mortality are subjected to children born to women of different genders, length of marriage, or time after first birth.

2.3.3 Three Different Methods of Indirect Estimation

Indirect estimation approaches utilize the indirect data from registered numbers of children dying/surviving and children ever born, or children dead proportion. Notably, Brass (1964) first developed the indirect estimation approaches in the 1960s. Number of literature works explained such approaches, specifically, a report published by United Nations in 1983 named as *Manual X: Indirect Techniques for Demographic Estimation*. This report specially mentioned such methods in its Chapter-III titled as Children Ever Born and Children Surviving (CEBCS) (Liu et al., 2015). Indirect estimation of childhood mortality requires to focus on the gathering of required data, challenges and limitations including three variations as:

1. Estimation of childhood mortality classified by age of the mother (AGE);
2. Estimation of childhood mortality classified by duration of marriage (DOM);
3. Estimation of childhood mortality classified by time since first birth (TSFB).

For several years, first two of such variations are widely taken into account as per the United Nations (1983). However, the third variation is the latest one and categorized on the basis of time since first birth.

Comparison of the three variants of the indirect method; Indirect methodology categorized in three key variations with their own benefits as well as limitations. Evaluating child mortality using indirect method is based on age of mother (AGE), duration of marriage (DOM) and time since first birth (TSFB).

(a) Age of mother (AGE)

The most commonly used variant depends on the age of women. However, some problems have been found while evaluating the information gathered from the youngest age groups, i.e. 20-24 or 15-19 years (5 years duration). Generally, these groups showed the over-estimation of childhood mortality rate. Such part depicts the actual impact of greater mortality in between the offspring born to the mothers who are younger age. However, other factors also influenced such as selection influence in which the women belongs from inferior socioeconomic groups tend to begin the childbearing at their younger ages as well as showed the greater mortality risks among their children (Alkema & You, 2012). Furthermore, the death and birth samples taken from the women who were younger is relatively lesser as needed. Therefore, there is greater chance of random errors in such groups. Several survey reports exacerbate such issues consisting of relatively small sample sizes and adjustments relatively low mortality as well as low fertility. Unfortunately, most recent points in time depends on the estimates provided by such age groups. The survey reports generally neglect such significant practice providing the evaluation depending on 15-19 age group as well as 20-24 age group.

Another issue happens related with the age-based variant. Survey report involves the data mainly about ever-married women (as per the results of several major organizations on ever-married samples). Therefore, there is the requirement for achieving the appropriate parity distribution aspects in this case. This factor expands the denominator of every age group (15-19 and 20-24) depending on the proportion inverse of women ever married belongs to every age group. According to the assumption, there is no variation in numerators representing the no births outside of marriage.

(b) Duration of Marriage (DOM)

The duration of marriage variant has been less influenced by the sample size as well as selection bias problems in between young women. It has the major benefit with respect to age variant as it offers superior evaluations on current child mortality. It can be used as another kind of selection bias. Although several people do not find it unusual to give birth outside of marriage, the mortality risk for children who have been born out of wedlock are not considered until the mother is married(if at all). In addition, at the period when the experience of such children is taken into account, it is geared towards the incorrect duration of exposure. For example, if a child was born six years ago, when the mother was 16, and the mother married at the age of 21 in the year before the survey, the exposure proxy (duration of marriage) does not represent more than one year of exposure when it was much longer in fact. This factor is significant if the ever married women have no adjustments in the duration of marriage. Otherwise, it adversely affects the results of the surveys as the incorrect samples. It is essential for never-married women too (where it can be assumed as they have no children) just opposite of age factor that needs these as an adjustment.

(c) Time since First Birth (TSFB)

Hill and Figueroa (1999) proposed more recently the variant as time since first birth. It prevents certain issues existing in the other two variants. Such as with respect to DOM variant, it has the merit over age variant, as it is not greatly influenced by small sample sizes as well as socioeconomic bias for the most current evaluations. TSFB also prevents the issues connected with significant birth numbering happening outside of marriage. TSFB does need the first birth date to be gathered. Generally, this factor is only available in the reports where surveys record complete full birth histories. MICS has the one exception as to enquire about first birth dates available in all reports together with the Brass questions. TSFB can also be simply utilized along with ever-married samples, just like the DOM variant.

Some other Limitations of Indirect Estimation

There is the high chance of failure of indirect approaches due to lacking of significant information. For instance, if women do not remember the age of their first child or even her own age. This issue is quite common in several low-developed or poor nations. The results also can highly be affected if the estimation of age is based on the classifications connected indirectly/directly with the levels of mortality; e.g., children numbers who ever take birth. The insufficient data in surveys/censuses or that are not properly made with respect to accurate data collection for mortality evaluation have also been illustrated as the suffering from blunder of reports regarding dead children. Conversely, some reports include the details about the cases consisting both live births as well as stillbirths as the answer to the question on the number of offspring ever born, therefore resulting as the misjudges of mortality rates (Neupert., et al., 2019).

In addition, the desecrations of assumptions are the most common problem occurred in case of indirect approaches to evaluate the child mortality. The indirect approaches with an implicit assumption representing the cohort birth by women i.e. the children took birth in specific time duration. There have been more factors of high risks in case of the births to women 20–24 (and in few reports to women 25–29) of mortality as compared to the children born within the last five years of a survey.

Indirect approach has another issue with the estimate location in time. Indirect approaches evaluate the probability of death depends on the experience that can expand over several years leads to the average over that duration. The approaches utilized to locate the mortality evaluate in time can be generated the less or more in error, based on variations in trends as well as fertility in child mortality. Due to several assumptions need to be taken into account, there is a requirement to utilize the indirect approaches as well as partial amount of information that is available.

Another weakness of the indirect approaches is that censuses and general surveys are not designed to collect data specifically for estimation of mortality hence have reports of dead children omitted (Uddin et al., 2009). Elsewhere, there are cases where stillbirths and live births have both been used to answer the question of the number of children ever born, resulting in the overestimation of mortality rates. This study uses the indirect estimation variant basing on the age of women to estimate under-five mortality. According to Hill and Figueroa (1999), there is high mortality experienced by children born to mothers in younger age groups such as 15-19 and 20-24. The woman's age variant has a limitation of overestimating the mortalities in the younger age groups.

2.4 Data Quality of Age Reporting

One of the essential demographic variables is age. Therefore its accuracy is of paramount importance in population studies. However, age is mostly misreported and, as a result, constitutes demographics frustrating problems due to reporting errors and irregularities, which impacts negatively on its usage (Denic et al., 2004). In 2012, a report by KNBs indicated that age is more susceptible to anomalies and is more misreported than sex (KNBS, 2012). The most common age data problems that have been documented are age overstatement, digit preference, and age heaping. Digit preference is more distinct in populations or sub-groups with low education status. These errors have been reported during age reporting in most Sub-Saharan countries (Yazdanparast et al., 2012). Though the patterns and causes differ from one society to the next, age preference for those ending with "0" and "5" is relatively widespread. Apparently, this has existed for much of human history (Pullum, 2006).

Studies on age preference and avoidance in developing countries have shown vast distortions (Yazdanparast et al., 2012; Pullum, 2006; West et al., 2005). This occurs mostly when age is unknown, so respondents or interviewers tend to estimate, leading to heaping on 0 or 5 (Pullum, 2006). Studies show that census in African countries usually suffers from digit preference, also referred to as a content (or response) error or non-random measurement error (ESCWA, 2013; West et al., 2005; Yazdanparast et al., 2012). Bocquier et al. (2011) noted a significant impact of age heaping on demographic, economic, and health statistics in Africa since the late 1960s and 1970s. Irregularities in age data from African and Asian samples have been noted by previous studies (Denic et al., 2004; Palamuleni, 2013), and recent work has examined age heaping in Nigeria and Zambia (Bello, 2016). However, the quality of census in terms of age-reporting has improved remarkably in Asia, but less so in African countries (Cleland, 1996). A study in Zambia shows that males were more inclined towards reporting digit "5" compared to females who preferred "0". Besides digits 0 and 5, preference for ages with digits ending with 2 and 8 in 1990, and 8 in 2010 by both sexes was also common, which agrees with other similar studies (Bello, 2016). In another study, data showed significant avoidance of ages ending with 1, 2, 3, 6, and 7 (Bwalya et al., 2015).

2.5 Summary of Operational Framework used

This study used a three-phase operational framework. Phase 1 entailed a preliminary study where the key concepts and background of the study were defined, and a sketch of the entire research methodology was established. The second phase entailed the theoretical study, where different pieces of literature were reviewed and associated with the research gap in line with the research objectives. This part involved relating different concepts, theories, and frameworks that are associated with childhood mortality. Finally, phase 3 involved the practical study that is the actual data analysis or processing, the establishment of findings, and discussion of the results after the study. This part also involved giving recommendations and drawing conclusions.

CHAPTER THREE: DATA AND METHODOLOGY

3.1 Introduction

This chapter defines the structure of the research procedure to be used. It includes sources of data, sampling procedure, data quality issues encountered and methods of data analysis that have been utilized to arrive at the study's results.

3.2 Source of Data

The data used in this study were secondary data from the 2015/2016 Kenya Integrated Household Budget Survey (KIHBS). Data utilized in the study included statistics on child deaths and childbirths in the 12 months between August 2016 and September 2015. The survey design deliberately captured a family range of health status indicators such as demographics, household income, expenditure and consumption patterns, among other indicators.

The survey interviewed 25,423 women who were aged 15-49 years and asked about lifetime fertility and mortality. The survey asked the following questions:

1. How old is (name)?
2. Have you ever had a live birth?
3. How many children have you borne alive?
4. How many children have you borne alive who usually live in the household?
5. How many children have you borne alive who usually live elsewhere?
6. How many children have you borne alive have died?

3.3 Data quality

Errors are prone to data of any kind, especially those arising from faulty respondent recall and history of births. Another compromise in the data quality is event omission where children who died after surviving for a few hours or days; this data is not always included in the survey figures. In some cultural setups, purposive underreporting is associated with emotional events. This error is related to respondents giving inaccurate information. The completeness of reporting deaths is investigated by using internal consistency to find out whether there was an underreporting of deaths. Errors resulting from respondent's recall can also result in misreporting of date of births and age at the death of children. In KIHBS, the maximum age for women is 49 years. This means that birth history data for previous periods are confined to mothers who are younger at the time of birth. Since the mother's age at the time of giving birth affects the survival of the child, mortality rates could be biased since only younger women are interviewed. Further, data on birth history is limited to the experience of children for mothers who are alive. The fact is that children of dead mothers may be subject to greater mortality risks. The more time in history dates back to a great extent, the greater is the proportion of high-risk children who are represented by data of the birth history.

To assess data quality, various methods have been developed of assessing age data quality. These include Myers' Blended Index, Whipples index, Age Ratio Scores (ARS) and Age Accuracy Index (AAI), united joint score accuracy index. This study used Whipples to assess age heaping, age ratios by sex and united joint score method to assess accuracy. This study used the age ratio method, Whipples index, and United Nations Joint score method to evaluate the accuracy of age data.

During data collection, errors in age reporting occur in many stages and adjustments and remedies required vary depending on the source and nature of the error. Interviewers' prior knowledge of age that they offer directly to the respondent and estimating the age when the respondent is not so certain about the exact age bring about errors in age.

3.3.1 Age Ratio Method

The quality of age-sex data is accurately assessed by the age ratio technique. It compares the age ratios of information under evaluation with respect to the standard age-sex ratio values. It estimates the displacement levels of reported age among the age groups. It represents the ratio of total population in a specific age group with respect to the multiplication with the third population age group with the sum of populations in that age group. Further, the percentage is calculated by multiplied with 100. Mathematical expression is given for five years age groups as follows:

$$Age\ ratio = \frac{P_a}{P_3(P_{a-5} + P_a + P_{a+5})} \times 100$$

Where P_a population of a particular age group; P_{a-5} is population in age group before; and P_{a+5} is population in age group after. In cases where data has no abnormal variations, population changes or irregularities are minimal and age ratios are assumed to be equal to 100. It represents the standard age ratio. Deviation of an age ratio from the standard is indicated by the average absolute deviation from 100. Lower age ratio devitions shows higher accuracy of the age data reported. Variations with respect to 100, excluding the reasonable external factors like calamities or migration that could impact certain age groups, represents the displacement or undercounting errors in the information.

3.3.2 Whipple's Index

This index determines age heaping or digits preference on either the last digits 0, 5 or even a combination of both digits in age ranging between 23 and 62. This index changes between 100, standing for no preference for zero or five, and five hundred demonstrating that only digits 5 and 0 were reported. This method excludes extreme old and early childhood age brackets that are affected by different reporting errors rather than age preference. The mean values of the Whipple's index vary between 100 and 500 based on the assumption that the total number of individuals either rises or declines linearly with an increase in age. The index is developed to determine the degree of preference for individuals with ages that end in 5 and 0. It is calculated in single years as a ratio of individuals aged between 25 and 60 with their ages ending in 5 or 0 as a proportion of $1/10^{th}$ or $1/5^{th}$ of the total number of individuals between 23 and 62 randomly (Kyei, 2018). In the event there is avoidance or dislike of ages ending in 5 or 0, then this index ranges from 0 to 100.

Heaping and concentration in terminal digits 0 and 5 are some of the causes of inaccuracy of age ratios using whipples index. It is only applicable for the reported ages within single years. **Procedure**: Terminal digits '0' and '5'

Step1: At numerator, consider the range of 23-62, sum all ages ending in terminal digits 0 and 5.

Step2: At denominator, sum the population in complete age range 23-62 inclusive.

Step 3: Express the percentage and sum of the numerator on the one-fifth of total of the denominator.

$$Whipple\ index\ for\ the\ 5 - year\ range = \frac{\sum \quad (P_{25} + P_{30} + P_{35} + \cdots + P_{60}) \times 100}{\frac{1}{5} \sum \quad (P_{23} + P_{24} + P_{25} + \cdots + P_{62})}$$

(a) Interpretation:

i. The variation in Whipple's index is from 0 to 500

ii. No reporting on the Index=0 digits ‘0’ and ‘5’
iii. no preference for ‘0’ or ‘5’ at Index =100,
iv. The age data reports the digits ‘0’ and ‘5’ only at Index=500
v. Highly accurate data when Index<105
vi. Fairly accurate data at Index 105-109.9
vii. Approximate data at Index 110-124.9
viii. Rough data at Index=125-174.5
ix. Very rough age data at Index>=175

Table 3. 1: Whipple Score for testing

Quality of data	Scale
Highly Accurate	less than 105
Fairly accurate	105 - 109.9
Approximate	110 – 124.9
Rough	125 – 174.9
Very Rough	175+

Source: Shyrock and Siegel (1976)

From the results, the Whipple's index for females who reported ages ending in zero is 97.98, while those who reported ages ending in five is 114.95. The data for ages ending in zero is highly accurate while that ending in five is approximated.

3.3.3 United Nations Joint Score Index

A joint accuracy index was created by the United Nations to summarize age and sex ratios. The five-year age consistency is measured using this age-sex accuracy index groups (United Nations, 1956). Five year age grouping eliminates errors within the age groups attributable to misreporting or age shifting. The index incorporates age-group data accuracy tests for both sexes to calculate the accuracy of the sex ratios of different age groups separately.

The UN age-sex accuracy index is then the sum of (a) average male age ratio variations of 100, (b) average female age ratio variations of 100, and (c) three times the average age variations of the sex ratios recorded. This index makes it possible to compare datasets by sex. This is because the age ratio scores and sex ratio scores are measured together. There are three classifications of the UN ranking: one is right, where the index is < 20; two is wrong, where the score is 20 and

40; and three is highly inaccurate, where the score is > 40. It should be noted that age-sex evolving from demographic changes should be carefully considered when analyzing the UN index, as the index is unable to identify inaccuracies resulting from abnormal changes.

Shryock et al. (1976) noted that the index's key drawback is that as fails to put the expected reduction in sex ratio due to age increase, age abnormalities due to epidemics, migration, and wars, and usual variations in deaths and births into account. Another drawback is that the index uses an age ratio that does not have an upward bias in the core age group. It should also be remembered that within the formula, a lot of weight is placed into the sex ratio (Shryock et al. 1976). The joint scores may be impacted by sex ratio differentials that favor women because males experience high mortality rates. Because of its use of 5-year age group data for males and females to measure age and sex data accuracy in a given population, the United Nations age-sex accuracy index was used in this analysis. The index uses age ratio and sex ratio scores to generate a merged score showing a population's age-sex data quality score.

3.4 Method Applied

The indirect method of calculating mortality rates was used to establish the levels of under-five mortality. This method operates under the assumption that mortality and fertility patterns have remained consistent in the period of the study. The study required three sets of data, including the number of women by age group, children ever born (CEB) and children dead (CD) by age of mother and children ever . From the 3 sets, this method transforms proportions of CD among the CEB to these women into conventional measures of mortality.

The parameters estimated from it are q1, q2, q3, q5, ql0, ql5 and q20. However, only q5 is considered in this study.

These estimates are given by women aged 15-49 based on their births and child deaths. Naturally existing is the bias arising from incomplete and data that is not fully representative. No data is available for dead women since only surviving women aged between 15-49 years are interviewed. There will be bias on the estimates of mortality in case a difference is seen between children born to surviving women and those born to women who were dead at the time of the survey. However, any method of estimating childhood mortality by mothers that are dependent on retrospective reporting is biased. Women above 40 years of age are left out of the survey and cannot contribute information on the exposure of deaths of their little ones for periods before the interview. This censoring of information and the resulting potential for bias becomes more severe as mortality estimates are made for periods more distant before the survey. To reduce the methodological limitations of child mortality in this report, a period of 15 years is restricted before the survey is conducted.

3.5 Data Analysis

The Brass indirect method of estimation, and in particular the Trussell variant for West model life tables, was utilized to determine the mortality rates for under-fives. The Brass method uses three main categories of information: the total female population by five year age group, the number of children ever born by age group of mother, and the number of children dead by five year age group of mother.

3.5.1 Nature of data Required

a) Children ever born and dead

The information on the number of children ever born and those that are dead are obtained in a census or survey by asking women in the age range of (15 to 49) on their experience in childbearing. The set of questions asked include:

The Brass technique employs the technique that data consisting of children ever born and dead being categorized by the mother’s age. The classification is done using the traditional five-year age groups which run from 15-19, to 45-49. The age-groups are utilized in purposes of tabulation hence the data can be used for estimation.

b) The female population of reproductive age

The reproductive age of females used for the estimation is in the range (15-49) irrespective of their marital status. Ages in a survey or a census can be provided by a third party and often the information from proxies is marred with errors.

3.5.1.1 Procedure of Computation

Step1. Calculating average parity per woman

Women in an age group of five years provide the average parity that indicates the mean number of children who have ever been borne. The mathematical expression is;

$P(i) = CEB(i) / FP(i)$

Where, *FP(i)* refers the total number of females in the age group in spite of their marital status, , *CEB(i)* indicates the children given birth by the women in the age group *i*, and *P(i)* represents the average parity of females in the age group i *i*. Likewise, such factor also involves the case in which women did not provide their answer to the question on children ever borne. The group inclusion includes the assumption about women without any children. There is the requirement of parity values the age groups 25-29, 20-24, and 15-19.

Although, the analysis on the quality of raw data based on the calculations of the rest of the age groups.

Step 2. Proportions of dead children

The proportion always indicate the ratio of children who are dead to the total number of children who are ever born.

$D(i) = CD(i) / CEB(i)$

Where, *D (i)* $D(i)$ indicates the proportion of children who are dead coming from the women categorized in the age group *i*, while *CEB (i)* refers the total number of children have ever borne from the women in age group i and *CD (i)* $CD(i)$ indicates those children who have died as registered by the women in age group *i*. $CEB(i)$

Step 3. Computing multipliers, k (i)

According to the Trussel version, the mathematical expressions to determine the multipliers is:

$q(x) = k(i) * D(i)$

Where,

$k(i) = a(i) + b(i) P(1)/ P(2) + c(i) P(2) / P(3)$

There is the relationship between dying probability exactly at age x as q(*x*)$q(x)$ connected with the proportion of dead children D (*i*) $D(i)$product with the factor $k(i)$k (i). The parity ratios

P(1)/P(2) and P(2)/P(3) and coefficient *a(i)* , *b(i)* and *c(i)* calculate the factors $\frac{P(1)}{P(2)}\frac{P(2)}{P(3)}$. The simulation method can provide the coefficients by using regression model.

Step 4. Measuring dying probabilities by age xx, *q(x)*

The separate measurements of $D(i)$ k (*i*) as well as D (*i*) $k(i)$ for each age group gives the evaluation of q(*x*) by obtaining the multiplication.

q (*x*) = D (*i*) * k (*i*)

Step 5. Measuring reference dates for $t(i)$ and $q(x)$.

Step 4 is used to attain each $q(x)$ leads to provide their respective reference time $t(i)$ during the situations such as variations in steady mortality with respect to the alteration in time. Such process can be expressed as the survey conducted in number of years. Coefficients utilized on parity ratios can evaluate this factor. Mathematically, reference time is calculated as:

$$t(i) = e(i) + f(t)\frac{P(1)}{P(2)} + g(i)\frac{P(2)}{P(3)}$$

The measured values of $t(i)$ can be altered in definite dates by obtaining the decimals that represents the values difference from the reference survey date.

Step 6. Transforming to a common index

Previously, the step number 4 and 5 provide the evaluation of $q(x)$ for ages of x of 1,2,3,5,10,15,20 as well as for $t(i)$ that represents the year count of past survey. Every evaluated value of $q(x)$ is transformed into a particular measure for assessing the trends and facilitating the comparisons between/within the information. The dying probability by age five $q(5)$indicates the purpose of suggested common index. Experts not recommend to utilize the infant mortality due to the sensitivity of evaluated $q(1)$ with respect to the mortality pattern underlying the changing models.

Needed conversions obtained by using the values of $q(x)$ with respect to the model-life-table family. Linear interpolation provides the actual conversion between the given values. Let assume that an evaluation of $q(x)$ represented by $q^e(x)$ is transformed to a $q^c(5)$ i.e. a corresponding value where $x \neq 5$. It is vital to get the mortality levels for a model-life-table family with respect to $q(x)$ values that are nearby the evaluated value, $q^e(x)$. Thus, j and $j+1$ can be detected as,

$$q^j(x) > q^e(x) > q^{j+1}(x)$$

Where $q^j(x)$ and $q^{j+1}(x)$ are the model values of $q(x)$ for levels j and $j+1$. Conversely, the estimated value is $q^e(x)$. So, the mathematical expression for desired common index is

$$q^c(5) = (1.0 - h)q^j(5) + hq^{j+1}(5)$$

Where the interpolation factor h is represented as,

$$h = \frac{q^e(x) - q^j(x)}{q^{j+1}(x) - q^j(x)}$$

The provided information included both the data on children who have died and the number of children ever borne for all sexes.

Step 7. Interpretation

It is important to plot the factors against time when we attain the seven estimates for every group i of the selected common index. Reference dates can be obtained from the $t(i)$ values via step 5. Afterwards, the plotting against date can be acquired from the chosen indexes $q^c(5)$.

3.5.2 Method of Estimating Mean Age of Childbearing

The mean at first birth defines the age as the mean value at which women have gone through their reproductive years and give birth to their first child. There is a significant variation in the population sizes of fertile age groups due to the effects of mortality/migration and time-varying sizes of birth cohorts. These asymmetrical age profiles may generate contradictory birth numbers and predicted populations when integrated with fertility curves of same total fertility but contrary age patterns. Mean age of childbearing is utilized to fit/model the fertility curves in order to accurate for population composition or cohort size variations.

Mean age evaluation of the women have their first birth are given for nations with accurately vital statistics. Although, Kenya i.e. a less-developed nation has insufficient vital information due to which the evaluations are not readily available. Thus, estimates are necessary in analyzing the childbearing commencement in low developed nations. Likewise, the trends utilize such estimates to present the onset period of child bearing. Mathematically, it is calculated as,

$$M^*(t) = \frac{\sum_0^{a_{max}} \quad p(a,t) - p(a_{max},t)a_{max}}{1 - p(a_{max},t)}$$

Where $M^*(t)$= the average childbearing age at time t

p_{max}= proportion of women never to have given birth at a_{max}

$p(a,t)$= The proportion of women not having borne a child at age a and time t

Data required to estimate MAC (mean age of child bearing)

Mean age at childbearing (MAC) represents the mean age of mothers at the birth of their children. There is an assumption as the measurements are related with women who are subject through age 50 to the age-specific fertility rates observed in a given year. It can be calculated as the total age-specific fertility rates weighted by mid-point age of each age group, divided by the sum of age-specific rates:

$$MAC = \frac{\sum_a \quad x_a f_a}{\sum_a \quad f_a}$$

where fa is the age-specific fertility rate for women in age group a and x_a represents the mid-point of each age interval (17.5, 22.5,…47.5.). The calculation was done using the formula as provided for in the Mortpak 4.3 which was used to run the under-five mortality rate.

3.5.2.1 Choice of the Model

The Coale–Demeny West model was used to calculate the under-five mortality since it was found to be the best "average" model compared to the North, South and East Models.

3.6 Assumptions

a) The rates of childhood deaths and fertility have not changed in the past.

b) Infant mortality and age patterns of fertility in the population are presented by model patterns utilized in establishing the method.

c) At no time will the children mortality fluctuate by 5 year age group of their mothers.

d) No relation exists between a mother's survival in a population and the mortality risk of children.

e) Any alterations in the death rates of children in the previous years must have been slow and without a specific direction.

CHAPTER FOUR: UNDER FIVE MORTALITY RATE

4.1. Introduction

This chapter presents the results of the under-five mortality rate. The chapter presents data quality, mean age of child bearing, differentials in proportion of children dead by women, and levels of under-five mortality at national and county level by the residence.

4.2 Data Quality

4.2.1 Age Ratios

Age ratios were done for both male and female in the forty seven counties to establish whether there was age misreporting in the 2015/16 KIHBS data. The findings are presented in Figure 4.1 below.

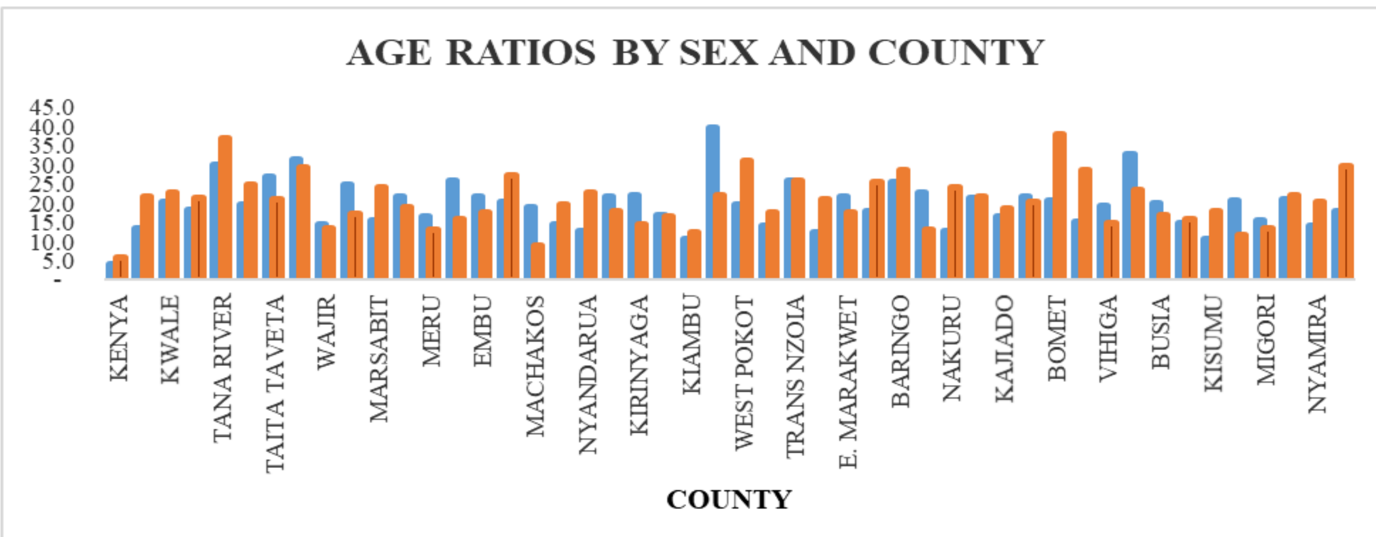

Figure 4. 1: Age Ratios by Sex and County

Figure 4.1 shows age ratios by sex and county. The average age ratio score was 4.1 for males and 5.9 for females. The data can, therefore, be considered highly inaccurate since it deviates highly from 100. An age ratio under or over 100 implies that persons were misclassified to an age group that is adjacent to their actual age. At the county level, high age ratio points for male were found in Tana River, Bungoma, and Turkana at 29.9, 32.7, and 39.7 respectively. For female age ratios, fluctuations are found in almost all age groups although with varying degrees of deviations. The counties with highly inaccurate age ratios for females were Tana River at 36.8, West Pokot at 31.2 and Nairobi at 29.6. Irregularities of these age ratios could be as a result of displacement from age groups. This could be attributed to interviewers misreporting age data in order to minimize their work of administering questionnaires to respondents.

4.2.2 Whipples Index

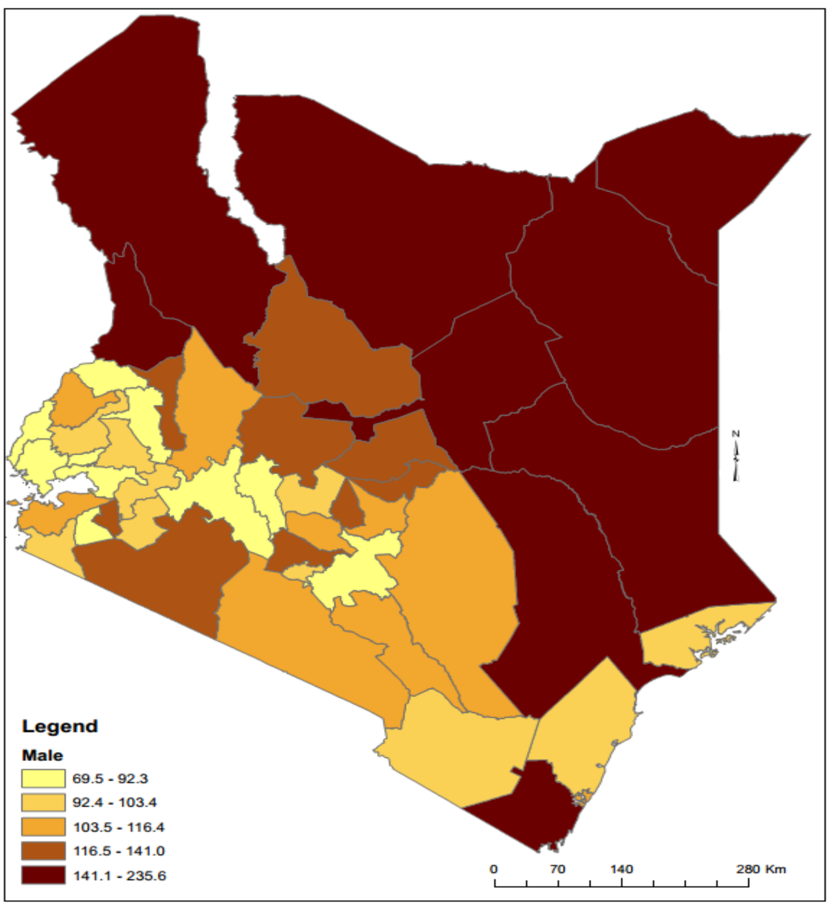

Figure 4.2: Whipples Index for Male

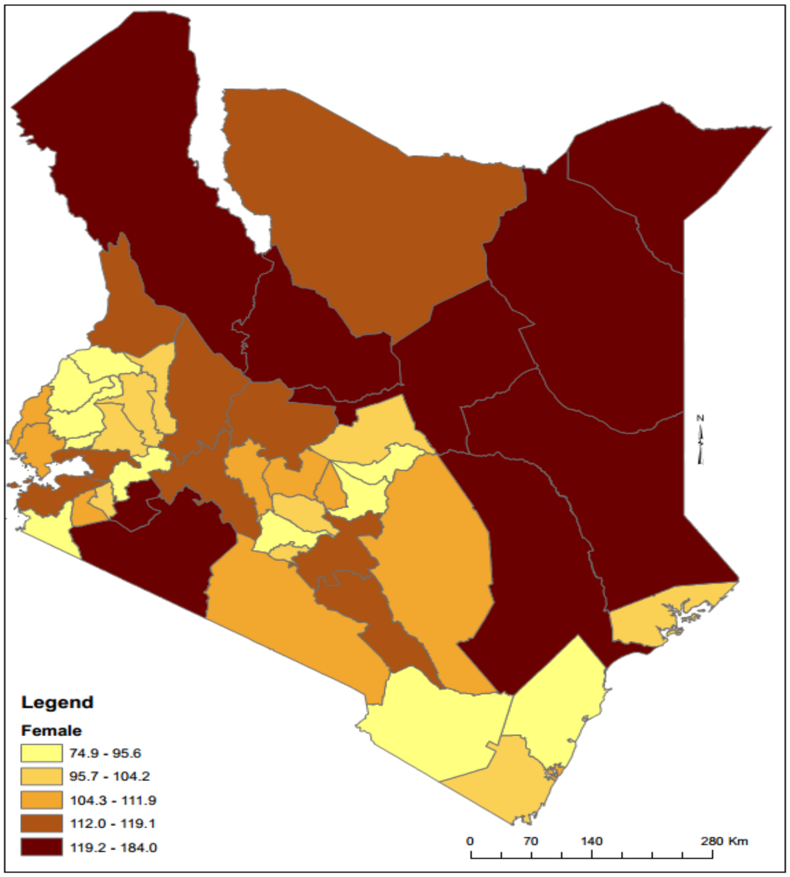

Figure 4.3: Whipples Index for Female

The counties with the highest whipple's index were Garrissa, Wajir, Mandera, and Isiolo each having an index of over 175. This shows that age distribution in these counties was very rough with age heaping at ages with terminal '0' and '5'. Kilifi, Lamu, Nyandarua, Nairobi, Migori, Vihiga, Kericho, Uasin Gishu, and Trans Nzoia counties had the lowest Whipple's indices (less than 105) for both males and females. These shows that age reporting by sex in these counties were highly accurate. The remaining counties had indices of at least 125 for either of the sexes showing that the data was rough. The findings show that age heaping is evident in forty three counties (See appendix 3).

In general, the results from age ratios, UN accuracy index and the whipples index show that KIHBs age data was of poor quality.

4.2.3 United Nations Joint Score

The United Nations joint score was used to determine whether the data from the forty seven counties were accurate, inaccurate or highly inaccurate. The findings show that the national accuracy score was 30.1 which is inaccurate. The county with the highest joint score is Baringo at 212.9 followed by Nairobi at 205.7 while Migori and Busia had the lowest scores of 79.5 and 88.2 respectively. All the forty-seven counties reported joint scores much higher than 40 indicating the 2015/16 KIHBS data is highly inaccurate.

Table 4. 1: United Nations Joint Score

County	Joint Score	County	Joint Score
Kenya	**30.1**		
BARINGO	212.9	KITUI	123.3
NAIROBI	205.7	VIHIGA	123.0
GARISSA	201.4	NYAMIRA	121.3
ISIOLO	192.1	NYANDARUA	121.2
BOMET	192.0	EMBU	118.1
WEST POKOT	191.9	THARAKANITHI	117.5
TANA RIVER	190.6	TURKANA	116.0
TRANS NZOIA	188.3	LAIKIPIA	115.9
KWALE	184.4	KISUMU	114.6
NAROK	178.0	MURANG'A	113.5
KILIFI	177.1	NYERI	112.3
MANDERA	175.7	KIRINYAGA	111.4
MARSABIT	175.1	MAKUENI	107.0
KAKAMEGA	166.9	KIAMBU	101.2
BUNGOMA	152.4	MACHAKOS	101.0
NANDI	151.5	SIAYA	100.8
KISII	140.7	SAMBURU	99.9
LAMU	137.0	UASIN GISHU	96.3
KAJIADO	136.3	KERICHO	95.6
NAKURU	135.5	MERU	94.0

WAJIR	133.8	HOMABAY	93.0
MOMBASA	131.1	BUSIA	88.2
E. MARAKWET	130.3	MIGORI	79.5
TAITA TAVETA	125.2		

4.3 Mean Age of Childbearing

The main aim of mean age at childbearing (MAC) was to establish the mean age of mothers at the birth of their children if subjected throughout their lives to the age-specific fertility rates. Considering the fact that 2015/2016 KIHBs data has no fertility data collected, KDHS 2014 data was used instead to estimate MAC. The findings are shown in table 4.3. The overall mean age of childbearing in Kenya is 28.14. The County with the lowest mean age of childbearing is Marsabit (25.96), followed by Nyamira (26.08) then Nairobi City (26.63). The counties with the highest MAC are Bungoma (29.12), Kilifi (29.04), Lamu (28. 89), and Garissa (28.83). Differentials in the MAC are normally consistent with the fertility patterns. However, the ASFR for age groups 45–49 may be slightly biased due to truncation. At times the 40- 44 age groups are also affected. Age specific fertility rates were women aged 20-24, 25-29, and 30-34 who indicated higher mean age of childbearing in all the fortyseven counties.

Table 4. 2: Mean Age of Childbearing

	Age Specific Fertility Rates(ASFR) per 1000							**Mean Age of**
COUNTY	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**Childbearing**
KENYA	**96**	**206**	**183**	**148**	**100**	**38**	**9**	**28.14**
Mombasa	81	162	163	118	79	41	..	28.08
Kwale	126	225	191	184	169	30	..	28.23
Kilifi	123	218	226	203	163	72	8	29.04
Tana River	144	300	221	202	127	133	..	28.68
Lamu	80	246	153	129	169	57	..	28.89
Taita/Taveta	47	187	137	134	111	27	..	28.71
Garissa	78	231	301	308	186	..	..	28.83
Wajir	129	355	323	399	182	..	..	28.04
Mandera	29	270	271	248	139	..	..	28.53
Marsabit	105	249	255	207	..	..	..	25.96

Isiolo	126	252	217	179	121	43	..	27.75
Meru	115	152	127	119	64	16	..	26.77
Tharaka-Nithi	63	187	138	110	90	52	..	28.54
Embu	61	204	157	94	82	5	9	27.36
Kitui	75	222	184	138	106	37	..	28.08
Machakos	79	186	163	107	85	36	20	28.39
Makueni	59	222	166	82	87	35	10	27.96
Nyandarua	48	219	189	129	94	21	..	27.96
Nyeri	52	181	154	81	52	17	0	27.04
Kirinyaga	52	130	130	80	57	19	..	27.68
Murang'a	21	191	196	88	74	35	0	28.39
Kiambu	47	133	156	128	70	14	..	28.26
Turkana	136	336	320	256	176	..	..	27.50
West Pokot	133	285	322	338	236	..	..	28.49
Samburu	170	282	260	251	172	107	..	28.68
Trans Nzoia	103	287	187	195	157	84	..	28.82
Uasin Gishu	108	179	178	120	76	57	..	27.83
Elgeyo Marakwet	70	244	223	156	101	23	..	27.76
Nandi	102	223	186	126	110	50	..	27.93
Baringo	76	272	234	156	112	..	..	27.24
Laikipia	96	182	217	109	96	32	..	27.66
Nakuru	69	205	181	144	112	36	..	28.39
Narok	225	303	240	230	120	83	..	27.36
Kajiado	118	172	207	202	112	63	..	28.68
Kericho	132	202	186	140	106	36	..	27.46
Bomet	130	244	178	187	85	42	..	27.38
Kakamega	101	240	237	161	110	23	8	27.73
Vihiga	83	217	241	196	117	38	..	28.40
Bungoma	103	256	180	205	188	68	0	29.12
Busia	128	269	227	167	101	49	..	27.45
Siaya	97	267	158	165	133	29	..	27.84

Kisumu	87	196	182	130	98	15	..	27.51
Homa Bay	178	285	239	201	111	29	..	26.87
Migori	136	317	257	189	111	46	..	27.31
Kisii	99	197	165	168	91	31	0	27.82
Nyamira	133	220	157	128	34	30	..	26.08
Nairobi	81	149	135	105	35	15	..	26.63

,, means data not available

Source: 2014 KDHS

4.4 Differentials in Proportion of Children Dead by mother's Age

The pattern of proportions dead of children ever born with maternal age was observed among the age groups per county. This was to check the quality of the data of children dead by age of the mother. The findings are shown in table 4.4.

At national level, the proportions dead of children ever born were higher for the 15-19-year age group than the 20-24-year age group. There was fluctuation in the proportion of children dead across different age groups at the county level. The highest proportion of children dead by women were recorded in women aged between 15-19 (0.901) followed by women aged between 45-49 (0.0808), 40-44 (0.0676), 35-39 (0.613) and lastly 25-29 (0.430) and 20-24 (0.424) had the lowest proportion of children dead by women. Overall, the proportion of children dead by women for those in age group 15-19 was higher than 20-24 in all the Counties except Tana River (0.1485), and Mandera (0.1969). The counties that recorded the highest average proportion of children dead by women were Migori, Mandera and Siaya at 0.1441, 0.1277 and 0.1191 respectively. Conversely, the counties that recorded the lowest were Laikipia, Meru and Nyeri at 0.0207, 0.0264 and 0.0276 respectively. The failure of the proportion dead to rise with age is an indication of poor quality data possibly due to under reporting of births by older women.

Table 4. 3: Differentials in Proportion of Children Dead By Age Group of Mother

	Age Group of Mother							
Region/ Residence	**15-19**	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**	**Average**
Kenya	0.0434	0.0297	0.0406	0.0472	0.0579	0.0677	0.0849	**0.0530**
Mombasa	0.0831	0.0296	0.0309	0.0414	0.1011	0.1509	0.0963	**0.0762**
Kwale	0.0684	0.0437	0.0560	0.0435	0.0743	0.0819	0.0742	**0.0631**
Kilifi	0.0581	0.0544	0.0549	0.0522	0.0827	0.0849	0.1270	**0.0735**
Tana River	0.0452	0.1485	0.0470	0.1288	0.0829	0.0924	0.1152	**0.0943**
Lamu	0.0750	0.0753	0.0676	0.0553	0.0881	0.0994	0.1568	**0.0882**
Taita Taveta	0.1623	0.0406	0.0471	0.1033	0.0821	0.0170	0.0720	**0.0749**
Garissa	0.0709	0.0432	0.0641	0.0496	0.0451	0.0978	0.0511	**0.0603**
Wajir	0.0687	0.0981	0.0460	0.1174	0.1000	0.1479	0.1451	**0.1033**
Mandera	0.1791	0.1969	0.0747	0.0692	0.1234	0.1754	0.0754	**0.1277**
Marsabit	0.0466	0.0069	0.0427	0.0365	0.0502	0.0630	0.0703	**0.0452**
Isiolo	0.0635	0.0524	0.0261	0.0048	0.0216	0.0075	0.1094	**0.0407**
Meru	0.0458	0.0070	0.0121	0.0079	0.0201	0.0188	0.0729	**0.0264**
Tharaka Nithi	0.0481	0.0359	0.1247	0.0482	0.1137	0.0413	0.1377	**0.0785**
Embu	0.0864	0.0737	0.0312	0.0249	0.0576	0.0159	0.0692	**0.0513**
Kitui	0.0470	0.0105	0.0432	0.0768	0.0560	0.0817	0.1214	**0.0624**
Machakos	0.1049	0.0101	0.0373	0.0336	0.0657	0.0392	0.0918	**0.0547**
Makueni	0.1199	0.0285	0.0072	0.0079	0.0463	0.0181	0.0448	**0.0390**
Nyandarua	0.1348	0.0479	0.0512	0.0752	0.0764	0.0113	0.0710	**0.0668**
Nyeri	0.0564	0.0437	0.0135	0.0248	0.0284	0.0046	0.0220	**0.0276**
Kirinyaga	0.1885	0.0439	0.0241	0.0190	0.0500	0.0474	0.0595	**0.0618**
Murang'a	0.2663	0.0662	0.0444	0.0552	0.0460	0.0523	0.0165	**0.0781**
Kiambu	0.3316	0.0417	0.0368	0.0313	0.0285	0.0179	0.0463	**0.0763**
Turkana	0.0338	0.0480	0.0411	0.0746	0.0480	0.0476	0.1321	**0.0608**
West Pokot	0.0809	0.0505	0.0530	0.0428	0.0566	0.0901	0.1067	**0.0686**

Samburu	0.0133	0.0154	0.0284	0.0024	0.0075	0.0372	0.0107	**0.0164**
TransNzoia	0.1348	0.0243	0.0802	0.0284	0.0548	0.0692	0.0376	**0.0613**
Uasin Gishu	0.0393	0.0283	0.0181	0.0157	0.0203	0.0533	0.0486	**0.0319**
Elgeyo Marakwet	0.1026	0.0122	0.0337	0.0416	0.0578	0.0659	0.0103	**0.0463**
Nandi	0.0857	0.0482	0.0212	0.0328	0.0428	0.0244	0.0435	**0.0427**
Baringo	0.0795	0.0152	0.0442	0.0083	0.0200	0.0427	0.0391	**0.0356**
Laikipia	0.0530	0.0269	0.0087	0.0119	0.0225	0.0178	0.0042	**0.0207**
Nakuru	0.0552	0.0447	0.0331	0.0190	0.0377	0.0796	0.0434	**0.0447**
Narok	0.0369	0.0250	0.0408	0.0249	0.0523	0.0627	0.0575	**0.0429**
Kajiado	0.0288	0.0191	0.0182	0.0281	0.0247	0.0600	0.0321	**0.0301**
Kericho	0.0569	0.0622	0.0329	0.0201	0.0492	0.0555	0.0368	**0.0448**
Bomet	0.0433	0.0031	0.0051	0.0331	0.0082	0.0300	0.0154	**0.0197**
Kakamega	0.0591	0.0479	0.0610	0.0560	0.0738	0.0655	0.1495	**0.0733**
Vihiga	0.1170	0.0588	0.0410	0.0579	0.1450	0.0520	0.1698	**0.0916**
Bungoma	0.0642	0.0301	0.0698	0.0634	0.0845	0.1268	0.1137	**0.0789**
Busia	0.0871	0.0202	0.0374	0.0614	0.0593	0.0894	0.1155	**0.0672**
Siaya	0.1508	0.0243	0.0818	0.1105	0.1329	0.1720	0.1617	**0.1191**
Kisumu	0.1023	0.0336	0.0358	0.0692	0.0487	0.1205	0.0979	**0.0726**
Homabay	0.0859	0.0442	0.0987	0.1031	0.1205	0.1434	0.1430	**0.1055**
Migori	0.0779	0.0209	0.1148	0.2027	0.1941	0.1914	0.2069	**0.1441**
Kisii	0.1009	0.0392	0.0255	0.0075	0.0152	0.0250	0.0692	**0.0404**
Nyamira	0.0399	0.0347	0.0058	0.0135	0.0386	0.0571	0.0564	**0.0351**
Nairobi City	0.1545	0.0173	0.0067	0.0113	0.0276	0.0303	0.0509	**0.0426**

4.5 Comparison of Under Five Mortality Rates with other Countries in Africa

The information presented in this section shows the probabilities of deaths by age five in different regions across Kenya. The results indicate a substantial difference between the countries and also among the forty seven counties in Kenya.

Table 4. 4: Under Five mortality Rate between Countries

Country	Survey	U5MR	Survey	U5MR	% Change
South Africa	2016 DHS	42	1998 DHS	59	28.8
DRC	2011-12 DHS	68	2005 DHS	117	41.9
Ethiopia	2016 DHS	67	2011DHS	88	23.9
Tanzania	2015-2016 DHS	67	2010	81	17.3
Uganda	2016 DHS	64	2011	90	28.9
Ghana	2014 DHS	60	2008	80	25.0
Kenya	2014 DHS	52	2008-09	74	29.7

Source: STATcompiler (2020)

All the countries had a decline in mortality rate. However, the U5MR reduced by 41.8% followed by Kenya at 29.8% while Tanzania had the lowest decline (17.3%).

Under Five Mortality Rate by County

The counties with the highest mortality rates per 1000 children were Migori (150), Mandera (131), and Wajir (106) while Meru (12), Samburu (17), Bomet (16), and Nairobi City (18) recorded the lowest under-five mortality rate. In general, the findings show that counties from Nyanza regions (Migori, Homabay, Siaya, and Kisumu), Western, Coastal (Tana River, Kilifi, Lamu, Taita-Taveta, and Kwale) and North Eastern (Mandera, Wajir, Garissa) region recorded high mortality rates compared to those from Central and Rift Valley.

Table 4. 5: Under Five Mortality Rate by County

COUNTY	U5MR (2015/2016) KIHBS
Kenya	52
Migori	150
Mandera	131
Wajir	106
Homa Bay	99
Siaya	98
Tana River	83

Lamu	83
Tharaka-Nithi	78
Vihiga	75
Bungoma	75
Mombasa	71
Kilifi	64
Kakamega	64
Taita/Taveta	60
Garissa	60
Kwale	59
Kisumu	59
West Pokot	58
Nyandarua	56
Murang'a	54
Turkana	54
Trans Nzoia	54
Busia	53
Kitui	52
Elgeyo/Marakwet	44
Kericho	44
Nakuru	43
Marsabit	41
Narok	40
Embu	38
Machakos	37
Kirinyaga	36
Nandi	36
Kiambu	34
Nyamira	30
Kajiado	29
Uasin Gishu	28

Nyeri	26
Baringo	26
Isiolo	25
Kisii	24
Makueni	23
Laikipia	18
Nairobi City	18
Samburu	17
Bomet	16
Meru	12

4.5.1 Checking Consistency in Under Five Mortality Rate with other Determinants of Mortality at Macro Level

To check if the calculated estimates are plausible, regression analysis was conducted to establish the relationship between under-five mortality rate in the forty seven counties and the women's fertility (TFR) and education (women with secondary education). The analysis involved six other countries including South Africa, Democratic Republic of Congo, Ethiopia, Uganda, Ghana, and Kenya. Results from the regression analysis were presented and discussed in connection to the literature review. T-value for each variable was computed to determine if the effect was significant or not. The regression model was as follows:

$$\mathbf{Y = \beta_0 + \beta_1 TFR + \beta_2 E_{it} + \varepsilon}$$

Y = Under-five mortality rate (U5MR)

TFR= Women fertility

E= Education level

B_0, B_1, and B_2= βi = regression coefficients (to be calculated)

e = error term

Table 4. 6: Summary of Regression Statistics

	Multiple R	R Square	Adjusted R Square	Standard Error	Observations
South Africa	0.45061	0.20305	-0.0246	11.9433	10
DRC	0.62176	0.38659	0.25027	19.5521	12
Ethiopia	0.77688	0.60354	0.51543	14.7084	12
Tanzania	0.27099	0.07343	0.02335	17.7	40
Uganda	0.72483	0.52538	0.48018	9.23064	24
Ghana	0.8900	0.7921	0.7690	9.5279	11
Kenya	0.44178	0.19517	0.15859	27.4238	47

Table 4.7 gives the model from where the equation that could fit the data was obtained. From the table, a positive correlation existed between the dependent and independent variables in all the countries. Ethiopia and Uganda had strong positive correlation between U5MR and the independent variable (Education and TFR). However, Tanzania and Kenya had weak correlation between the independent and dependent variables. DRC, Ethiopia, and Ghana had very strong correlation between independent and dependent variables. The adjusted R^2 was used to show the predictive power of the study model and it was found to be 0.79 in Ghana and 0.60 in Ethiopia implying that 79% and 60% of U5MR are explained by education and TFR. This indicated that the regression model fits the data well. For the case of South Africa, DRC, Tanzania, and Kenya, TFR and Education only explains 20%, 38%, 7%, and 19% variation between the independent and dependent variables. The smaller variation in these countries could be as a result of biasness in data and the smaller sample size used.

Table 4. 7: Combined ANOVA Results By Country

		df	*SS*	*MS*	*F*	*Significance F*
South Africa	Regression	2	254.41	127.20	0.89	0.45
	Residual	7	998.49	142.64		
	Total	9	1252.90			
DRC	Regression	2	2168.34	1084.17	2.84	0.11
	Residual	9	3440.58	382.29		

	Total	11	5608.92			
Ethiopia	Regression	2	2963.96	1481.98	6.85	0.02
	Residual	9	1947.04	216.34		
	Total	11	4911.00			
Tanzania	Regression	2	918.68	459.34	1.47	0.24
	Residual	37	11591.72	313.29		
	Total	39	12510.40			
Uganda	Regression	2	1980.66	990.33	11.62	0.00
	Residual	21	1789.30	85.20		
	Total	23	3769.96			
Ghana	Regression	2	3115.44	1557.72	16.94	0.00
	Residual	9	827.56	91.95		
	Total	11	3943.00			
Kenya	Regression	2	8024.63	4012.31	5.34	0.01
	Residual	44	33090.79	752.06		
	Total	46	41115.42			

The significance values for Ethiopia, Uganda, Ghana and Kenya were less than 0.05 ($\alpha<0.05$) which implies that the overall model for these countries was significant. Therefore, the relationship between U5MR and Education and TFR can be said to be statistically significant.

However, the relationship between U5MR and Education together with TFR was statistically insignificant in South Africa, DRC, and Tanzania as indicated by a significant level greater than 0.05.

4.5.2 Collinearity between Education and Total Fertility Rate

Collinearity test was done to check if highest education level and total fertility rate of women are correlated. The collinearity coefficient obtained by VIF value was 2.832, meaning that the VIF values are between 1 and 10. It can therefore be deduced that there is no collinearity between education level and TFR.

Table 4. 8: Collinearity Coefficient

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
	(Constant)	65.007	38.015		1.710	.094		
1	Education	-.684	.449	-.347	-1.523	.135	.353	2.832
	TFR	2.723	5.484	.113	.497	.622	.353	2.832

a. Dependent Variable: U5MR

4.5.3 Regression Coefficients by Country

The regression coefficient was run to establish the extent to which total fertility rate and women with highest level of education affects U5MR. The findings reveal that a unit increase in proportion of TFR of women is directly related to U5MR in South Africa, DRC, Tanzania, Uganda, Ghana, and Kenya. On the other hand, a unit increase in proportion of TFR of women is inversely related to U5MR in Ethiopia. With regards to the relationship between education and U5MR, a unit increase in proportion of women with higher educational level is inversely related to the level of U5MR in all the countries under study. However, the effect of both educational level and TFR are not statistically significant as the significant level in all cases was greater than 0.05 except for TFR in Ghana, and Uganda (P-value<0.05) and education level in Ethiopia (P-value <0.05). The findings show inconsistency which may be attributed by poor quality of data as there is no collinearity between the two independent variables.

Table 4. 9: Combined Regression Coefficients

Country	Variables	Coefficients	Standard Error	t Stat	P-value
South Africa	Education level	-1.4632	1.4462	-1.0118	0.3453
	TFR	16.1121	13.3415	1.2077	0.2664
DRC	Education level	-0.0735	0.6319	-0.1164	0.9099
	TFR	12.3681	8.5915	1.4396	0.1838
Ethiopia	Education level	-1.7535	0.6428	-2.7278	0.0233
	TFR	-6.9114	6.4508	-1.0714	0.3119

Tanzania	Education level	-0.2806	0.1644	-1.7072	0.0962
	TFR	0.1224	2.3230	0.0527	0.9583
Uganda	Education level	-0.0119	0.1345	-0.0887	0.9302
	TFR	10.2988	2.2993	4.4791	0.0002
Ghana	Education level	-0.0558	0.3024	-0.1846	0.8576
	TFR	15.8704	5.9158	2.6827	0.0251
Kenya	Education level	-0.6779	0.4491	-1.5094	0.1383
	TFR	2.7848	5.4868	0.5075	0.6143

4.5.3 County Level Output

The plausibility of the 2014 KIHBs U5MR data in the 47 counties was checked using regression residuals. The education of women at the county level was measured using the proportion of women with at least higher education. The counties with highest residual in absolute terms include Mombasa (30.37), Lamu (21.8), Wajir (25.27), Mandera (58.46), Marsabit (29.62), Isiolo (39.53), Meru (40.01), Tharaka-Nithi (26.92), Makueni (21.82), Samburu (54.13), Turkana (23.98), Laikipia (28.08), Bomet (37.67), Vihiga (27.26), Siaya (44.69), Homa Bay (41.50), Migori (86.93), West Pokot (17.02), Narok (21.63), Kakamega (13.87), Kajiado (13.37), Nandi (14.50), Uasin Gishu (11.30), Murang'a (14.43), Taita/Taveta (14.48), Garissa (13.57), and Nyandarua (10.54). The same counties exhibited higher UN joint score (>40), an indication of poor quality of data and inconsistency in the pattern of the proportion dead across the age groups, which are a result of misreporting or underreporting of children dead and surviving or violation of violation of assumptions in indirect estimation of childhood mortality.

Those with lower residuals are Kwale (5.24), Kilifi (2.91), Tana River (9.69), Embu (7.06), Kitui (5.73), Machakos (4.97) and Nyeri (6.01), Kirinyaga (7.20), Kiambu (3.41), Trans Nzoia (1.33), Elgeyo/Marakwet (3.46), Nakuru (4.90), Kericho (2.454), Busia (7.34), Nyamira (6.10), and Nairobi City (9.18). Despite the lower residuals in these counties, UN joint scores were highly inaccurate with unsteady pattern of the proportion of dead by women. This confirms that the 2015/16 KIHBs data are inconsistent in all the counties.

Table 4. 10: Regression Residuals

County	Estimated using indirect method	Predicted Y Using regression	Residual
Mombasa	71	40.50492471	30.36990303
Kwale	59	64.13647408	-5.246433238
Kilifi	64	60.8443323	2.914526942
Tana River	83	72.82590277	9.695134935
Lamu	83	60.92122108	21.80085928
Taita/Taveta	60	45.85996833	14.48305402
Garissa	60	73.32240604	-13.57852745
Wajir	106	80.42899749	25.27805552
Mandera	131	72.17182305	58.46838247
Marsabit	41	70.39473315	-29.62310751
Isiolo	25	64.76121283	-39.53060833
Meru	12	52.02110106	-40.01404875
Tharaka-Nithi	78	51.09411172	26.92078354
Embu	38	45.37813555	-7.068157055
Kitui	52	57.57039735	-5.735659209
Machakos	37	42.07865853	-4.970874836
Makueni	23	44.57938169	-21.81511853
Nyandarua	56	45.06854972	10.54776274

Nyeri	26	31.52058067	-6.01019905
Kirinyaga	36	42.74360984	-7.207893467
Murang'a	54	39.54125915	14.43878423
Kiambu	34	30.16487342	3.415275904
Turkana	54	77.8549219	-23.97621218
West Pokot	58	75.50443995	-17.02457306
Samburu	17	71.16794493	-54.12789253
Trans Nzoia	54	55.08991175	-1.335237214
Uasin Gishu	28	39.58527061	-11.30731376
Elgeyo/Marakwet	44	47.07833668	-3.469930539
Nandi	36	50.18912811	-14.50425273
Baringo	26	53.36593672	-27.68851983
Laikipia	18	46.30335672	-28.08061345
Nakuru	43	38.57582542	4.907776997
Narok	40	61.52041775	-21.63394904
Kajiado	29	42.22713157	-13.3728565
Kericho	44	46.66428926	-2.454067546
Bomet	16	53.39704586	-37.6709954
Kakamega	64	50.42182517	13.87378145
Vihiga	75	47.85331664	27.26005408
Bungoma	75	52.56718286	22.36872929
Busia	53	60.20492307	-7.346634486

Siaya	98	53.32192526	44.69208112
Kisumu	59	40.39869496	18.19862799
Homa Bay	99	57.8691116	41.50714944
Migori	150	63.29927583	86.93566991
Kisii	24	42.0328789	-18.00034733
Nyamira	30	36.39202334	-6.100052358
Nairobi City	18	27.18231748	-9.18231748

4.6 Discussion of Findings

This section discusses the results and how it relates with the previous findings. Given that KIHBS did not do any correctness on age reporting and children born and those dead after the survey, the choice of doing this study comes handy.

4.6.1 Data Quality

Data quality was assessed using UN scores, age ratios and whipples index. The UN joint score index revealed that all the data at the county level were highly inaccurate (>40). Age ratio results also showed that the deviation of age ratios from 100 was very high making them inaccurate. An age ratio under or over 100 could imply that persons were misclassified to an age group that is adjacent to their actual age. Irregularities of these age ratios could be as a result of displacement from age groups. This could be attributed to misreporting of ages by interviewers misreporting age data in order to minimize their work of administering questionnaires to respondents.

The study further established that all the counties had high whipples index, an indication that age distribution in these counties were very rough with age heaping at ages with terminal ‘0’ and ‘5’.

All the three data quality methods used show that the data was of poor quality in all the counties. Age tends to be more susceptible to anomalies and is more misreported than sex. Since it is a constant element in censuses and surveys, its misreporting constitutes one of the most demographic challenges. The quality issues in age reporting could be attributed by age overstatement, digit preference and age heaping.

As evident in whipples index, digit preference was more distinct at the county level. The findings are in line with studies done in other developing countries that age preference and avoidance have shown vast distortions (Spoorenberg, 2009; Pullum, 2006). This occurs mostly when age is not known, and so respondents or interviewers tend to estimate which leads to heaping on 0 or 5.

The inaccuracy in the 2015/2016 data is a continuation of errors made in these surveys. The 2003 DHS data was among the ones with more than ten percent of age displacement, especially for females across ages 15 and 50, at 16.1 percent displacement rate. The estimated level of women aged 40-45 were also misreported in the same data. Kenya is argued to be the fourth country in age misreporting with a percentage of 32.0 in the 2003 DHS data.

Age heaping was also witnessed in 2003 and 2009 data sets. The affected age group was 20-25 years. Wafula and Ikamari (2007) reports that 1998 and 2003 KDHs data sets revealed high underreporting of age amongst married women below 25 years, which suggested that there was age shifting from the other nearby ages. The same applied for ages 40 and 45. The 2009 DHS data showed digit preference for "0" and "5", and digit avoidance for odd numbers like 3, 7 and 9. The 2009 DHS data had less fluctuation despite the fact that there was under reporting at ages below 25 years.

4.6.2 Mean Age of Childbearing

The age-specific fertility rate was used to measure the number of births to women of a specified age group per 1,000 women in that age group. The results depict that Marsabit (25.96), Nyamira (26.08) and Nairobi City (26.63) had the lowest mean age of childbearing while Bungoma (29.12), Kilifi (29.04), Lamu (28. 89), and Garissa (28.83) had the highest. Women in age groups 20-24, 25-29, and 30-34 also indicated higher mean age of childbearing in all the forty seven counties than any other age group. However, the ASFR for age groups 45–49 and 40-44 may be slightly biased due to truncation.

The biasness in the older age groups may be associated with misreporting. In particular, it seems that the older age groups failed to provide their data so that their reported means remain smaller for most years than the means reported by younger age groups for the same years. In particular, the implications of misreporting the date of recent births influence fertility levels and trends. Schoumaker (2014) argues that such truncation errors in DHs surveys are non-trivial and lead to underestimation of total fertility rates (TFR). In addition to reporting errors in the birth history by mothers, most women misreport their own date of birth which could lead to biasness in mean age of childbearing data hence fertility data. Choumaker (2011) assessed the levels of birth omissions in DHS in Sub-Saharan Africa. The study showed that birth omission was a common error in all countries including Kenya, Guinea, Cameroon, Niger, Mali, Mozambique and Burkina Faso. These omissions introduced biases in fertility rates and indicated significant variations of fertility estimates between countries. Birth omissions and underreporting were more prevalent among recent births than longer period births. It showed that DHS questionnaires that were more complex and had longer reference periods produced significantly higher levels of omissions.

4.6.3 Differentials in Proportion of Children Dead By Women

The study found that proportions of children dead by mothers was high among mothers aged 15-19 years older and lowest for mothers aged 20-24. There is a steady rise in the proportion of children dead by mothers aged from 25-29 years to 45-49 years. The fact that there was no increase in the proportion of children dead by the age of women indicates that the age specific fertility data or ages of the mother were of poor quality. This could be as a result of older women not reporting their ages, and underreporting of new births. This is consistent with Mednick and Baker (2010) study that there is a linear relationship between maternal age and perinatal mortality with low rates among young women, and increasing rates with maternal age. Among the reasons for such high proportions of children dead among teenage mothers were; first, they are perceived to be physiologically immature; thus less desirable outcomes' another reason are differential social characteristics such as poverty, lack of access to prenatal care, ignorance, and poor nutrition.

In assessment of omission of births by Pullum (2014), the 2009 KDHS data revealed a report of less boys than were expected at birth with a deviation of -5.7 within 10 years prior to the survey. The highest level of incompleteness of date of birth was recorded by Guinea with 52.7 percent and Yemen with 45.9 percent. Omission of births was also found to be high in Dominican Republic and Armenia at 8.7 and 7.7 percent respectively. In Kenya, omission of births was at 3.6 percent.

4.6.4 Under Five Mortality Rates

Comparing the 2015/6 KIHBS data and 2009 census, there is a huge decline in the U5MR. The 2009 census showed Siaya (227), Kisumu (182) Migori (173), and homa Bay (170) as the counties with the highest mortality rates while Meru (48), Embu (49), Elgeyo/Marakwet (42), and Nandi (49) had the lowest U5MR. The change noticed is for a period of five years after the census. Overall, there was a huge decline in U5MR in all the counties with Kisii County recording the highest decline from 101 U5MR to 24 U5MR in 2009 and 2015/6 respectively.

As much as there is a decline in mortality rate in 2015/2016 data sets compared to the 2009 census, the results per county is not plausible. Comparing the U5MR with age data quality, it is worth noting that despite Migori (150), Mandera (131), and Wajir (106) recording high mortality rates while Meru (12), Samburu (17), Bomet (16), and Nairobi City (18) recording lower mortality rates in 2015/16 KIHBS data, age misreporting with high inaccuracies in all the 47 counties. Nairobi County recorded higher age heaping (102.3), and high inaccuracy as established by UN joint score results (205.7). The same high inaccuracies and age heaping were noted in Bomet (whipples index=131.7, UN joint score=192.0), Samburu (whipples index=149.0, UN joint score=99.9) and Meru (whipples index=119.5, UN joint score=94.0) counties.

Comparing the data quality of counties that recoded high U5MR, that is, Migori (150), Mandera (131), and Wajir (106). Migori had an average whipples index of 87 and accuracy score of 79.5 which shows that there was age heaping and misreporting in the County. The same case was witnessed in Mandera (whipples index=201, UN joint score=175.7), and Wajir (whipples index=206, UN joint score=133.8). It can be deduced that the mortality data collected at the county level depicts poor quality. With age and sex misreporting noted when assessing data quality, and mean age of child bearing, such biasness could result to inaccurate regression results in Kenya.

Regression results also show that a rise in proportion of TFR directly relates to under-five mortality rate while a rise in the number of mothers with highest education level inversely relates with U5MR. Twenty-seven counties including Mombasa, Lamu, Wajir, Mandera, Marsabit, Isiolo, Meru, Tharaka-Nithi, Makueni, Samburu, Turkana, Laikipia, Bomet, Vihiga, Siaya, Homa Bay, Migori, West Pokot, Narok, Kakamega, Kajiado, Nandi, Uasin Gishu, Murang'a, Taita/Taveta, Garissa, and Nyandarua all showed a poor fit in the relationship between estimated under five and education and fertility rates. The same counties exhibited higher inaccuracy score.

CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the summary of findings, conclusions, recommendations, and suggestion for future studies.

5.2 Summary of Findings

The quality of age data is important because age sex distribution is not only an invariable part of a survey report but the bias introduced in studies can lead to wrong inferences. The findings show that the age by sex, Age Specific Fertility Rates, proportion of children dead by mothers and under-five mortality rates were inaccurate as shown by age ratios, whipples index and UN joint score results. From this study, it emerged that the 2015/2016 KIHBS age data was inaccurately reported with large fluctuations in age ratios for males and females, which could be an indication of persons in various ages being carried across age group boundaries or persons misreporting their own ages for various reasons. This compromised the quality of data. The study found that ASFR for age groups 45–49 and 40-44 may be slightly biased due to truncation. However, ASFR was high for mothers aged 20's and 30's.

The proportions of children dead by mothers were found to be high among mothers aged 15-19 years older and lowest for mothers aged 20-24. His could be as a result of perceived physiologically immaturity, poverty, lack of access to prenatal care, ignorance, and poor nutrition among teenage mothers. The under-five mortality rate was also found not plausible in many counties.

5.2 Conclusion

Quality data is essential in producing accurate inferences in surveys. However, it is difficult to get high quality data, especially in Africa and Kenya in particular. The inaccurate results presented in the findings provided evidence that quality issues exist in the 2015/16 KIHBS data, characterized by inconsistency and inaccuracy of date of birth and current ages reported. The study therefore concludes that KIHBs data is not suitable to estimate under 5 mortality rate at county level because it did not have recent fertility (births in last 12 months), had inconsistent results for all the counties and the estimated U5MR is not related to either fertility of education.

5.3 Recommendations

Based on the above findings, the following policy and research recommendations are proposed.

5.3.1 Policy Recommendations

The study shows evidence of age misreporting that could have arisen due to reporter or interviewer bias. For respondents‟ bias to be addressed, agencies involved in the data collection need to also engage the public in understanding the importance of accurate reporting. This can be done by creating policies aligned to supporting mass awareness and educating people on how false reporting during surveys and censuses distorts information derived from data and how their lives are likely to be affected.

Since the results showed data anomalies from age misplacement and transfers, a more thorough supervision should be conducted during field work to ensure ages are recorded appropriately. This would be useful in reducing interviewer biases.

The analysis reveals evidence of age misreporting that may have happened because of the bias of the research assistants or interviewees (respondent). Research assistants need to be fully trained on collection of demographic indicators especially mortality in order to capture the information correctly and accurately. To address respondent bias, there is need to engage the members of the public on the importance of the survey before responding to the survey questions. This can be achieved by creating policies that can enhance mass awareness and how inaccuracies in the data can negatively affect their lives.

5.5 Recommendations for data collection on KIHBs

KIHBs should consider including a variable to measure recent fertility to help determine the mean age of childbearing. Partial birth history also requires further probes to remove under reporting.

REFERENCES

Alkema, L., & New, J. R. (2012). Progress toward global reduction in under-five mortality: a bootstrap analysis of uncertainty in Millennium Development Goal 4 estimates. *PLoS Med*, *9*(12), e1001355.

Bello, Y. (2016). Error Detection in Outpatients Age Data using Demographic Techniques. *International Journal of Pure and Applied Science and Technology*, *10*(1), 27 – 36.

Brass, W. (1964). *Uses of census or survey data for the estimation of vital rates*. UN.

Brass, W., & Ashley, J. COALE (1968): Method of Analysis and estimation. *The demography of tropical Africa, sous la direction de W. Brass et al., Princeton University Press, Princeton*, 88-139.

Bwalya, B. B., Phiri, M., & Mwansa, C. (2015). Digit preference and its implications on population projections in Zambia: Evidence from the census data. *International Journal of Current Advanced Research, 4*(5), 92-97.

Cleland, J. (1996). Demographic Data Collection in the Less Developed Countries. *Population Studies*, *50*, 433-450. Denic, S., Khatib, F., & Saadi, H. (2004). Quality of age data in patients from developing countries. *Journal of Public Health, 26*(2), 168-171.

Ferrarini, T., & Norström, T. (2010). Family policy, economic development, and infant mortality: a longitudinal comparative analysis. *International Journal of Social Welfare*, *19*. doi: 10.1111/j.1468-2397.2010.00736.

Heckmann, M. (2015). Determinants of the Recent Rise in Childhood Mortality in Sub-Saharan Africa and South Asia. *African Population Studies*, *28*(0), 1182. doi: 10.11564/28-0-566

Hill, K., and M-E. Figueroa (1999). Child Mortality Estimation by Time Since First Birth. Johns Hopkins Population Center Papers on Population WP 99-05. Baltimore, Maryland: Johns Hopkins Population Center. Available from http://www.jhsph.edu/popcenter/publications/pdf/WP99-05.pdf.

Hodge, A., & Jimenez-Soto, E. (2013). Determinants of childhood mortality.

Kenya National Bureau of Statistics (KNBS) and ICF Macro (2010) *Kenya Demographic and Health Survey 2014*. Calverton, Maryland: KNBS and ICF Macro.

Kenya National Bureau of Statistics (KNBS). (2012). *Kenya 2009 Population and Housing*

Census Analytical Report on Population Projections. Kenya National Bureau of Statistics, Ministry of Planning, National Development and Vision 2030, Nairobi.

Liu, L., Black, R. E., Cousens, S., Mathers, C., Lawn, J. E., & Hogan, D. R. (2015). Causes of child death: Comparison of MCEE and GBD 2013 estimates. *The Lancet, 385*(9986), 2461-2462. doi:10.1016/s0140-6736(15)61132-1

Mwangi D and Murrithi D(2015) "Determination of infant and child mortality in Kenya using Cox- proportional hazard model" American Journal of Theoretical and Applied Statistics 4(5):404-413.

Neupert, R., Menjivar, R. E. F., & Castilla, R. E. F. (2019). Indirect estimation of infant mortality in small areas. *Revista Brasileira de Estudos de População, 36*.

Okiro, E. A., & Ayieko, P. (2018). Childhood mortality during conflicts in Africa. *The Lancet*, *392*(10150), 804–805. doi: 10.1016/s0140-6736(18)31373-4

Pal, J., Mukhopadhyay, B., & Gupta, S. (2014). Myers' Blended Method: An Alternative Approach. *Journal of Scientific Research and Reports*, *3*(18), 2459–2465. doi: 10.9734/jsrr/2014/10673

Palamuleni, M.E. (2013). Age reporting in the North West Province, South Africa, 1996- 2007. Paper presented at the 2013, *Annual Meeting of the Population Association of America*, New Orleans, April 11- 13.

Pillay, T. (2019). Parent-Carer Education: Reducing the Risks for Neonatal and Infant Mortality. *Neonatal Medicine*. doi: 10.5772/intechopen.82786.

Pullum, T. W. (2006). An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003. *Maryland: Macro International Inc.*, 5. Calverton, Maryland: Macro International Inc.

Pullum, T. W., and Stan B. (2014). Evidence of Omission and Displacement in DHS Birth Histories. *DHS Methodological Reports No. 11*. Rockville, Maryland, USA: ICF International.

Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., Wang, H., Levin-Rector, A., Dwyer, L., ... & Murray, C. J. (2010). Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: a systematic analysis of progress towards Millennium Development Goal 4. *The Lancet*, *375*(9730), 1988-2008.

Renschler, J. P., Walters, K. M., Newton, P. N., & Laxminarayan, R. (2015). Estimated Under-

Five Deaths Associated with Poor-Quality Antimalarials in Sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene*, *92*(6_Suppl), 119–126. doi: 10.4269/ajtmh.14-0725.

Rutstein, S. O., and J. G. Rojas (2006). Guide to DHS Statistics. Calverton, Maryland: ORC Macro. Available from http://www.measuredhs.com/pubs/pdf/DHSG1/Guide_DHS_Statistics.pdf

Schoumaker, B. (2014). *Quality and consistency of DHS fertility estimates, 1990 to 2012*. Rockville, Maryland: ICF International.

Shryock H. S., Siegel J. S., & Associates (1976). *The Methods and Materials of Demography*. New York: Academic Press.

Spoorenberg, T. (2009). Is the Whipple"s index really a fair and reliable measure of the quality of age reporting? An analysis of 234 censuses from 145 countries. *New York: United Nations Department of Economic and Social Affairs*.

Sullivan, J. M., S. O. Rutstein, and G. T. Bicego (1994). Infant and child mortality. DHS Comparative Studies, No. 15. Calverton, Maryland: Macro International Inc. Available from http://www.measuredhs.com/pubs/pub_details.cfm?ID=23

Uddin, M., Hossain, M., & Ullah, M. O. (2009). Child Mortality in a Developing Country: A Statistical Analysis. Journal of Applied Quantitative Methods, 4(3), 270-283.

United Nations (1983). Indirect Techniques for Demographic Estimation, Manual X. *Population Studies, No. 81*. Sales No. E83.XIII.2.

United Nations Department of Public Information. (2015). *End poverty 2015, Millennium development goals, high-level event on the Millennium Development Goals, United Nations Headquarters, New York, 25 September 2008 = Éliminer la pauvreté 2015, objectifs du Millénaire pour le développement, réunion de haut niveau sur les objectifs du Millénaire pour le développement, siège des Nations Unies, New York, 25 septembre 2008*. New York (N.Y.).

United Nations Inter-agency Group for Child Mortality Estimation (UNIGME) (2018). Still Births and Child Mortality Estimates. https://childmortality.org/

United Nations. (1956). *Manuals on Methods of Estimating Population. Manual II: Methods of Appraisal of Quality of Basic Data for Population Estimates*. Sales No. E.56. XIII.2.

Wafula,S., & Ikamari, L. (2007). Patterns, levels and trendsin unmet need for contraception: a case study of Kenya. *https://www.researchgate.net/publication/281937670*

West, K.K., Robinson, J.G., & Bentley, M. (2005). Did Proxy Respondents Cause Age Heaping in the Census 2000? *National Academies of Sciences, Washington, DC 3*(6), 58–65.

Yazdanparast, A., Pourhoseingholi, M.A., & Abadi, A. (2012). Digit preference in Iranian age data. *Iranian Journal of Public Health. 9*(1).

APPENDICES

Appendix 1: Age Groups

AGE GROUPS	CHILDREN EVER BORN	CHILDREN DEAD	No of WOMEN
KENYA			
15-19	315,926	13,699	2,417,214
20-24	1,950,244	57,947	2,138,099
25-29	4,108,585	166,827	1,950,645
30-34	4,790,231	225,927	1,561,788
35-39	5,148,226	297,838	1,295,231
40-44	4,381,635	296,695	995,412
45-49	3,556,590	301,910	714,914
MOMBASA			
15-19	7,570	629	43,654
20-24	49,786	1,476	79,959
25-29	117,997	3,649	74,599
30-34	110,339	4,569	55,338
35-39	114,225	11,547	39,592
40-44	86,952	13,125	25,907
45-49	88,772	8,546	22,978
KWALE			
15-19	7,340	502	34,775
20-24	43,664	1,907	34,912
25-29	89,061	4,991	35,249
30-34	110,294	4,798	28,532
35-39	108,976	8,095	22,932
40-44	74,645	6,113	15,847
45-49	37,446	2,779	7,097
KILIFI			
15-19	7,498	436	67,488
20-24	78,671	4,280	74,455
25-29	105,252	5,780	53,314
30-34	216,780	11,319	61,821
35-39	175,182	14,496	34,014
40-44	113,588	9,640	22,188
45-49	70,997	9,017	16,403
TANA RIVER			
15-19	4,746	215	19,095

20-24	12,313	1,828	9,576
25-29	35,335	1,659	13,319
30-34	36,827	4,744	8,208
35-39	29,003	2,404	6,278
40-44	42,137	3,894	6,407
45-49	16,693	1,924	2,503
LAMU			
15-19	1,201	90	7,308
20-24	3,460	261	3,869
25-29	15,637	1,057	5,662
30-34	13,169	729	4,349
35-39	13,489	1,188	3,064
40-44	16,341	1,623	3,460
45-49	9,941	1,559	2,135
TAITA TAVETA			
15-19	3,243	526	20,314
20-24	9,730	395	13,662
25-29	16,783	791	10,073
30-34	34,240	3,536	11,600
35-39	34,435	2,828	11,399
40-44	39,977	680	10,446
45-49	33,630	2,421	7,340
GARISSA			
15-19	1,988	141	25,406
20-24	10,503	454	9,618
25-29	28,356	1,819	11,540
30-34	87,442	4,334	18,775
35-39	53,604	2,417	9,131
40-44	64,615	6,318	9,375
45-49	37,059	1,893	6,323
WAJIR			
15-19	1,557	107	20,698
20-24	20,589	2,019	14,623
25-29	50,080	2,303	13,198
30-34	63,299	7,429	11,184
35-39	74,702	7,472	11,195
40-44	59,505	8,804	7,168
45-49	40,935	5,939	5,600
MANDERA			
15-19	3,142	563	29,135

20-24	42,535	8,377	23,166
25-29	79,933	5,974	23,568
30-34	102,739	7,113	18,508
35-39	100,560	12,411	13,189
40-44	66,425	11,648	8,647
45-49	46,103	3,475	6,076
MARSABIT			
15-19	1,373	64	16,832
20-24	7,827	54	7,597
25-29	41,827	1,784	11,931
30-34	41,672	1,522	9,426
35-39	31,024	1,557	6,048
40-44	36,635	2,310	6,285
45-49	22,874	1,608	3,472
ISIOLO			
15-19	346	22	9,762
20-24	6,989	366	6,788
25-29	15,272	399	6,678
30-34	20,058	96	6,571
35-39	17,441	377	3,550
40-44	14,275	106	2,936
45-49	10,696	1,170	2,095
MERU			
15-19	8,253	378	77,219
20-24	58,310	406	56,683
25-29	81,091	979	47,837
30-34	122,520	964	50,772
35-39	160,025	3,215	51,421
40-44	133,624	2,511	40,979
45-49	153,346	11,177	32,141
THARAKA NITHI			
15-19	2,058	99	19,128
20-24	10,925	392	14,378
25-29	26,809	3,342	11,707
30-34	38,742	1,866	12,856
35-39	41,055	4,668	12,655
40-44	53,629	2,214	12,893
45-49	30,155	4,153	6,601
EMBU			
15-19	1,620	140	31,360

20-24	24,733	1,822	19,669
25-29	32,387	1,010	21,163
30-34	44,870	1,118	17,044
35-39	56,706	3,264	18,469
40-44	68,543	1,087	16,724
45-49	48,533	3,358	12,556
KITUI			
15-19	5,295	249	61,536
20-24	37,154	388	39,088
25-29	95,546	4,123	44,104
30-34	126,230	9,692	33,721
35-39	116,228	6,509	29,720
40-44	159,810	13,057	29,365
45-49	103,937	12,621	18,103
MACHAKOS			
15-19	1,297	136	67,593
20-24	41,153	418	62,967
25-29	99,868	3,728	54,519
30-34	75,285	2,526	30,035
35-39	104,272	6,849	31,982
40-44	108,753	4,265	30,157
45-49	139,033	12,768	30,544
MAKUENI			
15-19	4,253	510	62,232
20-24	24,088	687	35,289
25-29	75,181	538	36,455
30-34	76,297	600	24,286
35-39	102,967	4,766	28,787
40-44	100,265	1,819	22,325
45-49	85,460	3,831	18,734
NYANDARUA			
15-19	3,375	455	36,989
20-24	13,080	626	18,824
25-29	40,259	2,063	22,438
30-34	91,074	6,851	26,886
35-39	72,977	5,574	18,877
40-44	79,807	902	21,212
45-49	66,381	4,710	14,669
NYERI			
15-19	1,896	107	36,854

20-24	14,538	635	26,850
25-29	52,726	713	35,192
30-34	70,307	1,744	32,101
35-39	79,264	2,248	29,153
40-44	83,940	385	28,857
45-49	100,068	2,204	27,156
KIRINYAGA			
15-19	5,184	977	36,787
20-24	23,230	1,019	25,579
25-29	29,503	712	18,786
30-34	51,828	987	23,494
35-39	66,264	3,316	25,184
40-44	54,638	2,589	19,401
45-49	47,458	2,823	14,227
MURANG'A			
15-19	3,362	895	53,388
20-24	26,603	1,761	32,085
25-29	69,218	3,073	41,660
30-34	75,872	4,187	28,039
35-39	112,868	5,192	32,264
40-44	124,986	6,540	33,884
45-49	108,851	1,791	30,059
KIAMBU			
15-19	8,282	2,746	95,942
20-24	48,892	2,040	86,977
25-29	133,796	4,920	92,234
30-34	212,881	6,662	95,460
35-39	232,354	6,621	81,936
40-44	189,526	3,385	58,854
45-49	81,273	3,761	22,900
TURKANA			
15-19	17,747	600	46,233
20-24	54,118	2,598	41,821
25-29	80,468	3,310	26,956
30-34	153,421	11,452	34,999
35-39	139,543	6,695	28,209
40-44	156,821	7,471	27,694
45-49	124,156	16,401	21,308
WEST POKOT			
15-19	6,655	539	36,984

20-24	41,998	2,119	28,536
25-29	70,353	3,727	22,066
30-34	61,772	2,645	11,011
35-39	85,442	4,835	13,999
40-44	81,537	7,344	11,196
45-49	70,751	7,549	8,974
SAMBURU			
15-19	2,222	30	13,125
20-24	23,987	370	13,236
25-29	26,635	757	8,659
30-34	38,424	93	8,580
35-39	30,516	228	5,560
40-44	36,969	1,375	5,737
45-49	18,834	202	3,257
TRANSNZOIA			
15-19	6,867	926	50,163
20-24	51,264	1,244	57,742
25-29	92,087	7,382	35,723
30-34	80,727	2,295	28,487
35-39	89,044	4,883	21,847
40-44	136,172	9,424	22,211
45-49	80,096	3,012	13,648
UASIN GISHU			
15-19	9,032	355	69,559
20-24	31,927	902	57,719
25-29	96,937	1,758	58,387
30-34	111,593	1,750	37,850
35-39	159,063	3,228	42,963
40-44	77,862	4,148	20,235
45-49	71,621	3,481	13,183
ELGEYO MARAKWET			
15-19	4,553	467	29,044
20-24	19,760	240	23,613
25-29	47,986	1,617	18,007
30-34	36,928	1,536	10,800
35-39	68,783	3,977	13,777
40-44	65,847	4,337	10,994
45-49	34,669	357	6,687
NANDI			
15-19	5,966	511	64,628

20-24	40,231	1,939	47,675
25-29	79,591	1,687	39,456
30-34	112,234	3,680	30,410
35-39	92,765	3,972	24,349
40-44	89,611	2,185	21,239
45-49	50,709	2,208	9,114
BARINGO			
15-19	4,101	326	38,372
20-24	30,669	466	28,848
25-29	68,875	3,041	30,476
30-34	64,902	537	17,773
35-39	55,479	1,112	12,417
40-44	76,855	3,278	16,219
45-49	68,453	2,676	11,570
LAIKIPIA			
15-19	2,491	132	27,563
20-24	16,192	436	19,824
25-29	38,042	331	19,542
30-34	66,796	795	17,659
35-39	64,701	1,455	16,802
40-44	57,730	1,026	14,540
45-49	42,426	177	9,428
NAKURU			
15-19	8,869	490	90,606
20-24	91,783	4,103	100,099
25-29	176,774	5,859	77,881
30-34	213,398	4,045	77,503
35-39	279,228	10,521	69,664
40-44	219,454	17,476	53,060
45-49	108,168	4,690	22,593
NAROK			
15-19	13,031	481	57,153
20-24	88,937	2,226	57,782
25-29	127,113	5,190	40,776
30-34	109,087	2,718	26,100
35-39	126,168	6,602	23,159
40-44	124,652	7,812	22,640
45-49	89,073	5,124	12,533
KAJIADO			
15-19	5,971	172	46,817

20-24	39,155	748	47,318
25-29	95,101	1,733	55,801
30-34	102,122	2,868	39,209
35-39	80,635	1,992	21,437
40-44	59,545	3,573	15,782
45-49	48,798	1,566	13,943
KERICHO			
15-19	5,834	332	51,818
20-24	45,440	2,825	43,449
25-29	85,467	2,808	43,018
30-34	117,377	2,360	35,286
35-39	102,972	5,069	25,411
40-44	85,751	4,762	18,951
45-49	99,719	3,670	18,900
BOMET			
15-19	7,161	310	44,522
20-24	36,124	113	35,172
25-29	116,300	587	45,454
30-34	91,822	3,039	24,499
35-39	117,640	966	23,446
40-44	100,703	3,024	19,441
45-49	66,184	1,019	9,412
KAKAMEGA			
15-19	15,859	937	123,681
20-24	74,244	3,554	80,884
25-29	194,178	11,844	71,783
30-34	155,912	8,734	44,061
35-39	209,221	15,441	44,053
40-44	192,368	12,609	34,001
45-49	169,388	25,325	29,353
VIHIGA			
15-19	2,079	243	34,465
20-24	11,675	686	18,693
25-29	40,135	1,646	19,957
30-34	61,883	3,582	18,748
35-39	78,700	11,409	17,530
40-44	44,941	2,339	10,202
45-49	76,997	13,071	14,911
BUNGOMA			
15-19	25,802	1,656	101,493

20-24	63,722	1,919	49,886
25-29	196,199	13,692	64,287
30-34	136,299	8,637	36,843
35-39	181,720	15,356	38,162
40-44	130,821	16,587	22,590
45-49	224,951	25,580	35,138
BUSIA			
15-19	7,980	695	54,505
20-24	48,888	989	43,272
25-29	60,245	2,255	21,621
30-34	92,478	5,675	24,021
35-39	89,529	5,306	18,564
40-44	77,905	6,964	12,437
45-49	80,183	9,260	13,598
SIAYA			
15-19	7,031	1,060	63,838
20-24	54,393	1,322	38,518
25-29	90,512	7,403	31,616
30-34	113,404	12,532	27,177
35-39	119,285	15,852	23,742
40-44	84,424	14,519	17,582
45-49	83,190	13,450	14,406
KISUMU			
15-19	8,971	918	58,343
20-24	65,504	2,200	56,487
25-29	113,515	4,061	49,167
30-34	123,744	8,566	44,063
35-39	102,750	5,003	25,471
40-44	79,206	9,546	17,607
45-49	57,809	5,659	10,937
HOMABAY			
15-19	13,307	1,143	71,819
20-24	70,171	3,103	41,324
25-29	131,594	12,994	36,960
30-34	147,766	15,228	30,656
35-39	137,477	16,566	25,649
40-44	137,895	19,769	22,993
45-49	90,284	12,908	14,542
MIGORI			
15-19	28,615	2,230	85,814

20-24	55,042	1,153	45,604
25-29	125,613	14,416	38,976
30-34	158,525	32,135	33,430
35-39	111,261	21,592	19,932
40-44	112,231	21,483	17,407
45-49	121,970	25,230	17,740
KISII			
15-19	13,950	1,407	80,882
20-24	44,734	1,754	56,714
25-29	121,253	3,089	59,499
30-34	120,221	901	36,619
35-39	185,710	2,827	41,228
40-44	133,027	3,323	30,165
45-49	75,734	5,241	15,369
NYAMIRA			
15-19	6,349	253	39,562
20-24	22,603	783	24,179
25-29	63,181	366	26,234
30-34	81,998	1,108	27,711
35-39	88,431	3,409	21,597
40-44	68,248	3,894	18,293
45-49	65,034	3,671	14,020
NAIROBI CITY			
15-19	8,176	1,263	162,730
20-24	218,912	3,777	383,090
25-29	438,513	2,940	323,117
30-34	414,635	4,671	229,287
35-39	454,542	12,527	155,423
40-44	178,446	5,409	78,879
45-49	157,749	8,029	50,638

Appendix 2: Age Ratio

COUNTY	Age Ratio	
	Male	Female
Kenya	4.1	5.9
Mombasa	13.6	21.9
Kwale	20.4	22.6
Kilifi	18.1	21.5
Tana River	29.9	36.8
Lamu	19.5	25.0
Taita Taveta	26.8	20.9
Garissa	31.5	29.3
Wajir	14.5	13.5
Mandera	24.7	17.2
Marsabit	15.6	24.1
Isiolo	21.8	18.9
Meru	16.5	13.2
Tharakanithi	25.9	16.0
Embu	21.8	17.7
Kitui	20.4	27.2
Machakos	18.9	8.9
Makueni	14.5	19.5
Nyandarua	12.9	22.9
Nyeri	21.6	17.9
Kirinyaga	22.1	14.5
Murang'a	16.9	16.7
Kiambu	10.5	12.5
Turkana	39.7	22.2
West Pokot	19.6	31.2
Samburu	14.0	17.6
Trans Nzoia	25.9	26.0

Uasin Gishu	12.4	21.0
E. Marakwet	21.8	17.6
Nandi	17.8	25.7
Baringo	25.7	28.8
Laikipia	22.7	13.1
Nakuru	12.7	24.2
Narok	21.5	21.8
Kajiado	16.7	18.6
Kericho	21.7	20.4
Bomet	20.6	37.9
Kakamega	15.2	28.7
Vihiga	19.2	14.8
Bungoma	32.7	23.6
Busia	20.0	17.0
Siaya	14.9	15.8
Kisumu	10.7	17.9
Homabay	20.7	11.7
Migori	15.4	13.5
Kisii	21.2	22.2
Nyamira	14.0	20.3
Nairobi	17.8	29.6

Appendix 3: Whipples Index by Gender per County

County	Male	Female	County	Male	Female
Mombasa	107.3	109.2	Samburu	149.0	139.4
Kwale	144.5	101.4	Trans Nzoia	83.6	95.6
Kilifi	96.2	93.4	Uasin Gishu	71.6	99.1
Tana River	170.2	149.1	Elgeyo Marakwet	122.7	103.2
Lamu	100.1	100.5	Nandi	101.5	103.7
Taita/Taveta	100.6	85.4	Baringo	106.7	117.4
Garissa	208.0	181.8	Laikipia	117.5	119.1
Wajir	235.6	178.0	Nakuru	90.3	113.1
Mandera	219.7	184.0	Narok	126.2	126.8
Marsabit	141.2	118.6	Kajiado	111.3	106.5
Isiolo	171.2	163.4	Kericho	97.7	94.8
Meru	119.5	104.2	Bomet	103.4	131.7
Tharaka-Nithi	124.0	74.9	Kakamega	95.1	81.0
Embu	109.5	95.0	Vihiga	84.4	88.1
Kitui	130.5	111.3	Bungoma	104.1	95.6
Machakos	69.5	115.7	Busia	91.2	111.9
Makueni	105.2	114.3	Siaya	73.8	107.7
Nyandarua	86.3	104.5	Kisumu	92.0	115.7
Nyeri	98.5	106.5	Homa Bay	113.5	113.2
Kirinyaga	129.1	109.8	Migori	97.6	78.3
Murang'a	116.4	102.9	Kisii	78.5	110.3
Kiambu	116.6	92.5	Nyamira	117.1	101.1
Turkana	152.9	146.4	Nairobi	102.3	96.1
West Pokot	143.1	112.4			

Appendix 4: 2016 DHS Survey of South Africa

Characteristic	**Education Level**	**Total fertility rate 15-49**	**Under-five mortality rate**
Total	88.9	2.6	51
Western Cape	84.6	2.1	43
Eastern Cape	92.4	2.9	64
Northern Cape	83.7	2.7	51
Free State	85.8	2.4	63
KwaZulu Natal	90.1	2.5	44
North West	87.8	3.1	65
Gauteng	91.2	2.6	46
Mpumalanga	86.7	3	70
Northern Province	91.3	3.1	34

Appendix 5: 2013-14 DHS Survey of DRC

	Education Level	TFR	U5MR
Total	47.7	6.6	112
Kinshasa	88.9	4.2	83
Bas-Congo	49.6	6	124
Bandundu	50.8	6.3	89
Equateur	35.5	7	132
Orientale	34.4	5.9	112
Nord-Kivu	42.5	6.5	65
Maniema	41.8	6.9	105
Sud-Kivu	36.5	7.7	139
Katanga	42.2	7.8	121
Kasaï Oriental	46.8	7.3	122
Kasaï Occident	38.3	8.2	135

Appendix 6: 2016 DHS Survey of Ethiopia

	Education Level	TFR	U5MR
Total	17.2	4.6	82
Tigray	24.8	4.7	59
Afar	7	5.5	125
Amhara	17.9	3.7	85
Oromia	12.1	5.4	79
Somali	6.6	7.2	94
Benishangul-Gumuz	15.8	4.4	98
SNNPR	13.4	4.4	88
Gambela	34.5	3.5	88
Harari	29.3	4.1	72
Addis Ababa	54.3	1.8	39
Dire Dawa	28.5	3.1	93

Appendix 7: 2015-16 DHS Survey of Tanzania

	Education Level	TFR	U5MR
Total	23.4	5.2	78
Western	12.1	6.7	69
Tabora	16.4	6.6	72
Shinyanga	35.2	6.1	117
Kigoma	11.5	6.9	66
Northern	15	4.2	56
Kilimanjaro	12.9	3.4	50
Tanga	29.4	4.4	64
Arusha	37.8	4.6	51
Manyara	26.2	6	39
Central	28	5.7	66
Dodoma	19.7	5.2	98
Singida	13.2	6.2	50
Southern Highlands	17.9	4.3	65
Iringa	23.5	4.4	73
Rukwa	32.3	6.4	106
Lake	13.1	6.4	88
Kagera	17.5	5.5	88
Mwanza	14.3	6	88
Mara	23.4	6.7	94
Eastern	21	3.9	85
Dar es Salaam	44.1	3.3	94
Pwani	19	4.6	90
Morogoro	20.7	4.9	66
Southern	13.5	3.8	79
Lindi	13.4	4	79
Mtwara	13.5	3.6	79
Ruvuma	19.5	4.4	67
Zanzibar	66	5.1	56
Njombe	19.8	4	54
Simiyu	16.8	7.5	63
Geita	10.8	7.1	80
South West Highlands	20.1	5.2	95
Mbeya (before 2016)	24.1	4.5	95
Katavi	9.6	6.9	75
Kaskazini Unguja	53.5	6.5	62
Kusini Unguja	71.7	5.7	50
Mjini Magharibi	77.5	3.6	56

Kaskazini Pemba	43.2	7.2	57
Kusini Pemba	56	6.6	52

Appendix 8: 2016 DHS Survey of Uganda

	Education Level	TFR	U5MR
Total	32.9	5.4	73
South Buganda	17.2	4.7	59
North Buganda	52.9	5.4	74
Kampala	40.3	3.5	64
East Central	70	6.1	84
Eastern	38.3	5.9	65
West Nile	27.4	6	86
Northern	16.4	5.3	68
Karamoja	5.4	7.9	102
Western	24.4	5.7	84
South West	27.4	4.8	70
Busoga	38.3	6.1	84
Bukedi	24.6	6.1	72
Bugisu	32.2	5.6	68
Teso	26.4	6	54
Lango	13.6	5.1	68
Acholi	21	5.5	69
Bunyoro	22.9	6	89
Tooro	25.6	5.4	81
Ankole	27.4	4.9	72
Kigezi	27.2	4.6	67
Islands	26.1	6.2	96
Mountains	28.7	5.5	67
Greater Kampala	68.9	3.6	47

Appendix 9: 2014 DHS Survey of Ghana

Characteristic	Education Level	Total fertility rate 15-49	Under-five mortality rate
Total	63.1	4.2	70
Western	65.9	3.6	56
Central	67	4.7	69
Greater Accra	77.5	2.8	47
Volta	58.7	4.3	61
Eastern	68	4.2	68
Ashanti	73.8	4.2	80
Brong-Ahafo	57	4.8	57
Northern, Upper West, Upper East	27	5.9	99
Northern	23.1	6.6	111
Upper West	31.3	5.2	92
Upper East	32.9	4.9	72

Appendix 10: Education Level, Fertility, and U5MR Data

COUNTY	Education Level	TFR(KDHS)	U5MR(Computed KIHBS)
Mombasa	48.5	3.2	71
Kwale	19.8	4.7	59
Kilifi	26.3	5.1	64
Tana River	11.5	5.8	83
Lamu	22.9	4.3	83
Taita/Taveta	40.6	3.2	60
Garissa	12	6.1	60
Wajir	8.5	7.8	106
Mandera	10	5.2	131
Marsabit	11.8	5.0	41
Isiolo	19.7	4.9	25
Meru	31.1	3.1	12
Tharaka-Nithi	33.7	3.4	78
Embu	40.9	3.1	38
Kitui	26.2	3.9	52
Machakos	47	3.4	37
Makueni	42.9	3.3	23
Nyandarua	43	3.5	56
Nyeri	59.7	2.7	26
Kirinyaga	41.5	2.3	36
Murang'a	49.1	3.0	54
Kiambu	61.7	2.7	34
Turkana	8.6	6.9	54
West Pokot	13.3	7.2	58
Samburu	16	6.3	17
Trans Nzoia	35.2	5.2	54
Uasin Gishu	51.5	3.6	28
Elgeyo/Marakwet	42.5	4.1	44
Nandi	37.5	4.0	36
Baringo	36.1	4.8	26
Laikipia	42	3.7	18
Nakuru	53.4	3.7	43
Narok	29	6.0	40
Kajiado	51.3	4.5	29
Kericho	42.7	4.0	44
Bomet	34	4.3	16

Kakamega	38.8	4.4	64
Vihiga	43	4.5	75
Bungoma	38.1	5.0	75
Busia	25.6	4.7	53
Siaya	33.7	4.2	98
Kisumu	50.3	3.6	59
Homa Bay	31.1	5.2	99
Migori	23.5	5.3	150
Kisii	48.3	3.7	24
Nyamira	55.8	3.5	30
Nairobi City	66.1	2.7	18