

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Impact of HIV on estimates of child mortality derived using the summary birth history (CEB/CS) method

A dissertation submitted to
the Faculty of Commerce of the University of Cape Town
in partial fulfilment of the requirements for
the Degree of Master of Philosophy in Demography

By
Themba Mutemaringa

Centre for Actuarial Research

November 2011

PLAGIARISM DECLARATION

I, Themba Mutemaringa, declare that this research is original and was produced with assistance from my supervisor. I have cited and referenced all contributions from other people using the Harvard Convention. This research has not previously been submitted for examination purposes at this or any other university.

Signed: _____

Date: _____

University of Cape Town

ACKNOWLEDGEMENT

I would like to thank my supervisor, Professor Rob Dorrington, for the invaluable guidance during the research and throughout the course. I appreciate his patience and the time he spent, even after hours, in providing assistance and prompt feedback.

I also appreciate the work of Professor Tom Moultrie and the entire staff of the Centre for Actuarial Research and the Hewlett and Mellon Foundation for providing the scholarship which has facilitated this study and research.

Last but not least, I would like to glorify the name of the Lord for giving me the strength. I acknowledge and appreciate the priceless support I received from my family and friends.

University of Cape Town

ABSTRACT

The main objective of this study is to estimate the extent of bias introduced by HIV into the estimates of infant and under-five mortality derived from the Brass children ever born children surviving (CEB/CS) method.

The bias is estimated by comparing the infant and under-five mortality derived from the CEB/CS method with direct estimates from the full birth history data of infant and under-five mortality corrected for underestimation due to HIV, derived from recent Demographic and Health Survey (DHS) data. The correction factors used to adjust the infant and under-five mortality estimates obtained from the DHS data were derived using the method used by the Inter-Agency Group for child Mortality Estimation (IGME).

Estimates of overall bias of up to 15% in infant and under-five mortality rates derived from the CEB/CS method using data from women aged 25-39 were observed in all the six countries studied. Estimates of bias from data derived from women aged 20-24 showed different patterns in different countries. The results from these younger women could be affected by the intrinsic differences between the CEB/CS method and direct estimates from full birth history data. In two of the six countries, estimates of overall bias of more than 30% were observed. It was also observed that the bulk of the overall bias is due to the effect of HIV on the survival of mothers and their children. Estimates of bias resulting from the use of model life tables which included the impact of HIV were also calculated. The results showed that the choice of model life table does not introduce much bias, especially in under-five mortality estimates where five of the six countries showed biases in the range of +/-3%. It is recommended that further research should be done in other countries and into ways to correct for non-reporting of infant and under-five mortality resulting from selection bias due to HIV-related deaths of mothers before surveys.

TABLE OF CONTENTS

PLAGIARISM DECLARATION	2
ACKNOWLEDGEMENT.....	3
ABSTRACT.....	4
TABLE OF CONTENTS	5
LIST OF TABLES.....	7
LIST OF FIGURES.....	8
1 INTRODUCTION.....	9
1.1 Background.....	9
1.2 Main objectives of the research.....	11
1.3 Specific objectives	12
1.4 Significance of the study	12
1.5 Structure of the thesis	12
2 LITERATURE REVIEW.....	14
2.1 Introduction.....	14
2.2 Child mortality data collection systems.....	14
2.3 Measurement of child mortality	16
2.4 Child mortality trends	16
2.5 Approaches to estimating child mortality	23
2.6 Relative bias in methods of estimating childhood mortality.....	27
2.7 Effect of HIV on the methods of estimating child mortality.....	28
2.8 Methods for correcting for HIV bias on child mortality estimates	29
3 DATA SOURCES AND METHODS	35
3.1 Data Sources	35
3.2 Data Quality	35
3.3 Methods for estimating infant and child mortality.....	38
3.4 Indirect estimation of child mortality	39
3.5 Methods used to adjust infant and under-five mortality rates for underestimation due to HIV.....	40
3.6 Estimating bias in mortality estimates derived from the CEB/CS method.....	41

4	RESULTS	45
4.1	Estimation of infant and under-five mortality before correcting for the impact of HIV in Zimbabwe	45
4.2	Comparison of trends from direct and indirect estimates	48
4.3	Estimation of bias in survey estimates of childhood mortality	49
4.4	Comparison between estimates of infant and under-five mortality corrected using the Hallett <i>et al</i> and Hill and Walker methods.....	51
4.5	Estimation of overall bias in estimates derived using the CEB/CS method.....	53
4.6	Comparison of bias estimates among countries with generalised HIV epidemics.....	56
5	DISCUSSION AND CONCLUSION.....	60
5.1	Introduction	60
5.2	Limitations of the study	62
5.3	Scope for future research.....	63
5.4	Conclusion	64
	REFERENCES	65

LIST OF TABLES

Table 3.1	Births transference indices by country and DHS.....	38
Table 4.1	Direct childhood mortality estimates for Zimbabwe.....	46
Table 4.2	Indirect childhood mortality estimates for Zimbabwe	47
Table 4.3	Estimates of under-reporting (%) from the Hallett <i>et al.</i> and Hill and Walker models for five years preceding the survey.....	50
Table 4.4	Correction factors from the Hill and Walker model.....	50
Table 4.5	Comparison of estimates of infant and under-five mortality corrected using correction factors from Hallett <i>et al.</i> and Hill and Walker methods	52
Table 4.6	Overall bias in infant and under-five mortality estimates derived from the CEB/CS method for the 2005-06 ZDHS.....	54
Table 4.7	HIV non-survivor and regression equation bias, ZDHS 2005-06	55
Table 4.8	Estimates of model life table conversion bias for Zimbabwe	56
Table 4.9	Overall bias (percent) by age group of women and country	57
Table 4.10	HIV non-survivor and regression equation bias (percent) by age group and country.....	58
Table 4.11	Conversion bias (percent) in under-five mortality by age group of women and country.....	59
Table C.5.1	Conversion bias in IMR by country and age group of women	76
Table E.5.2	DHS samples sizes by country and survey.....	80

LIST OF FIGURES

Figure 2.1	Under-5 mortality estimates from various sources: Lesotho	19
Figure 2.2	Under-5 mortality estimates from various sources: Kenya	19
Figure 2.3	Under-5 mortality estimates from various sources: Malawi	20
Figure 2.4	Under-5 mortality estimates from various sources: Namibia	21
Figure 2.5	Under-5 mortality estimates from various sources: Zambia	22
Figure 2.6	Under-5 mortality estimates from various sources: Zimbabwe	22
Figure 4.1	Infant and under-five mortality from direct estimation; Zimbabwe 1988, 1994, 1999 and 2005 DHSs	47
Figure 4.2	Infant mortality and under-five from the CEB/CS method; Zimbabwe 1988, 1994, 1999 and 2005-6 DHSs	48
Figure 4.3	Direct and indirect estimates of under-five mortality for Zimbabwe.....	49
Figure 4.4	Ratios of estimates of infant and under-five mortality corrected using correction factors from Hallett <i>et al.</i> , Hill and Walker methods to the DHS estimates	53
Figure 4.5	Comparison of biases in estimates of infant and under-five mortality derived using the CEB/CS method.....	56
Figure A.5.1	Comparison of CEB/CS estimates and corrected IMR and U5MR for Kenya ...	68
Figure A.5.2	Comparison of CEB/CS estimates and corrected IMR and U5MR for Lesotho .	68
Figure B.5.3	Comparison of CEB/CS estimates and corrected IMR and U5MR for Kenya ...	71
Figure B.5.4	Comparison of CEB/CS estimates and corrected IMR and U5MR for Lesotho .	71

1 INTRODUCTION

1.1 Background

The level of child mortality is one of many important indicators used in determining the general level of development in a population. As such, concerted efforts should be made to establish accurate levels of child mortality to ensure that policy is properly informed. With such blueprints like the UN Millennium Development Goals, number 4 in particular, the need for correct measurement of child mortality cannot be over-emphasised. Over the past decades, child mortality in sub-Saharan Africa has been declining owing mainly to advances in health care. However, the advent of HIV just over two decades ago has seen a reversal in the gains made in reducing child mortality (Mahy, 2003). It is not surprising that there has been much interest in research about the association of HIV and child mortality, in particular about child mortality trends and estimation of the portion of child mortality attributable to HIV (Marston *et al.*, 2005; Timaeus, 1998; Adetunji, 2000; Ng'weshemi *et al.*, 2003; Garenne and Gakusi, 2006). Although recently, research has been done assessing the HIV-induced bias and fine-tuning the methods of estimating child mortality, much of the work has focused on direct methods of estimating child mortality (Hallett *et al.*, 2010; Artzrouni and Zaba, 2003; UNICEF, 2010; Inter-agency Group for Child Mortality Estimation, 2010b). Few studies have focussed on assessing the impact of HIV on indirect methods, especially the Brass CEB/CS method used for estimating child mortality (Ward and Zaba, 2008).

Unfortunately, most countries that have been affected by HIV are known to have problems with data. Sub-Saharan countries with more than two-thirds of known HIV cases have the most incomplete vital registration systems. This means that measures such as child mortality rates cannot, reliably, be computed directly using these data. Household surveys, such as Demographic and Health Surveys, offer the best alternative for estimation of child mortality, but they are expensive and require expertise to conduct and are affected by recall errors. Indirect methods represent an attempt to overcome the difficulties that arise in the collection of information about deaths of children especially those that happened a considerable time ago. These methods are based on data that can be collected from censuses or single-round surveys and the questions necessary to collect such data are quite simple and straightforward.

One of the challenges posed by HIV in countries with a generalised epidemic is its effect on conventional indirect methods of estimating child mortality. The CEB/CS method is widely used to estimate child mortality in developing countries where data

from vital registration are inadequate. The CEB/CS technique uses reports of summary birth histories of women obtained from surveys to compute the proportions of children born to women in five-year age groups that have died. The proportions of children dead are used to estimate standard life table probabilities of dying in a population, which in turn are used together with an appropriate model life table to produce a time series of estimates of child mortality.

HIV affects the CEB/CS technique by violating assumptions that are necessary for its implementation (Ward and Zaba, 2008). The CEB/CS method assumes that the mortality of children is independent of that of their mothers. HIV-positive mothers may pass on the virus to their children by means of vertical transmission, thus rendering invalid, the assumption of no correlation between mortality of mothers and their children. In breastfeeding populations, in the absence of a PMTCT programme, estimates of the proportion of children born to HIV-positive mothers who become infected range from 25 to 45 per cent (Zaba *et al.*, 2003).

HIV threatens the validity of the CEB/CS technique by introducing heterogeneity in child mortality with respect to the age of the mother, thus violating the assumption that survivorship of the children of women in a particular age group is representative of mortality of children of all women in the population. HIV prevalence varies by age, implying that the mortality of children of women in particular age groups may not be representative since the proportion of children born HIV-positive may not be consistent across the whole population. HIV prevalence also varies with time, as more people get infected the prevalence levels rise and variations may also be observed in response to interventions.

HIV also alters the age-specific mortality patterns resulting in inaccurate probabilities of dying before a given age since these probabilities are obtained through the use of multipliers that depend on model life tables. The model life tables used in the CEB/CS method were developed from non-AIDS populations so the use of such life tables in computing child mortality estimates lead to inaccurate results.

Child mortality estimates, obtained by applying the CEB/CS technique, from countries with generalised HIV epidemics, are subject to bias that could significantly distort the estimates of child mortality. The bias arises from the fact that most children born to HIV-positive mothers are at a higher risk of dying directly due to HIV or indirectly from problems associated with the death of their mothers. Mothers of these children are less likely to be included in surveys to report on the deaths of their children

because they either die between the time of giving birth and the time of the survey or because of illness (Hallett *et al.*, 2010).

This bias, if not allowed for, results in an underestimation of child mortality. Previous efforts to develop HIV correction factors by Ward and Zaba (2008) were based on the assumption of a stable epidemic, which does not hold since incidence levels have not been constant in most countries due to the natural progression of the epidemic and intervention strategies.

Hallett *et al* (2010) and the Inter-Agency Group for Child Mortality Estimation (2010b) each developed models that can be used to estimate the extent of bias in direct estimates of infant and child mortality resulting from the HIV-related correlation between the mortality of mothers and their children using fertility history data from surveys. Hallett *et al* (2010) used childhood mortality data collected from prospective household surveys in Manicaland in Zimbabwe to estimate, empirically, the bias in direct childhood mortality estimates resulting from HIV-induced correlation between deaths of mothers and their children. A mathematical model was then developed to estimate the bias in direct estimates of national infant and under-five mortality using cross-sectional survey data from Zimbabwe and other African countries.

The Inter-Agency Group for Child Mortality Estimation model (referred in this study as the Hill and Walker model) produces estimates of bias that can be used to adjust child mortality estimates from survey data in populations affected by HIV (Inter-agency Group for Child Mortality Estimation, 2010b). The model uses annual births, HIV prevalence data from pregnant women aged 15 to 49, mortality risks for HIV-negative children obtained from Coale-Demeny West life table and mortality risks of HIV-positive children obtained from cohort studies to derive the bias estimates.

The current research will use survey estimates of infant and under-five mortality adjusted for HIV-related bias using the correction factors derived from the Hallett *et al.* and Hill and Walker models to estimate the extent of bias in infant and under-five mortality estimates derived from the CEB/CS method.

1.2 Main objectives of the research

The main objective of this research is to estimate the magnitude of bias introduced by HIV to child mortality estimates derived from the children ever born/children surviving method by comparing them against direct estimates corrected for the impact of HIV.

1.3 Specific objectives

1. To adjust the direct estimates of infant and under-five mortality for underestimation due to the impact of HIV using the correction factors developed by Hallett *et al* and Hill and Walker for the following countries: Kenya, Lesotho, Malawi, Namibia, Zambia and Zimbabwe.
2. To compute estimates of child mortality rates from DHS data using the CEB/CS method for the same countries, without adjusting for the violation of assumptions underlying the method due to HIV.
3. To determine the extent of bias by comparing estimates obtained from the use of the CEB/CS method with the corrected direct estimates.
4. To establish and quantify the components of overall bias in mortality estimates derived from the summary birth history method.

1.4 Significance of the study

This study was prompted by the need to adjust estimates produced by the CEB/CS method for the effect of HIV/AIDS, since data for the implementation of the method are readily available especially from the current round of censuses. The method is currently being applied by most Central Statistics Offices without adjustment regardless of the suspicion about the validity of the method in the wake of HIV. Correcting child mortality estimates for bias will help determine reliable levels that allow for comparisons of child mortality estimates among countries and within a single country for different periods. Reliable trends will enable the correct evaluation and assessment of the effectiveness of programmes aimed at reducing child mortality as well as in assessing progress made towards achieving millennium development goal number 4 (Hallett *et al.*, 2010).

1.5 Structure of the thesis

The research comprises five chapters. Chapter 2 reviews the systems used in collecting child mortality data, child mortality trends and methods for child mortality estimation. Chapter 3 presents sources of data used, assessment of the quality of these data, methods for deriving correction factors used to adjust infant and under-five mortality estimates derived from the DHS data for underestimation, as well as methods for estimating the various types of bias in estimates derived from the CEB/CS method. The estimates of infant and child mortality derived from both direct and indirect methods, corrected infant and under-five mortality rates and estimates of bias are presented in

Chapter 4. Chapter 5 concludes the thesis with a discussion of the results and an assessment of the findings in the context of the theoretical framework for child mortality estimation and suggest ideas for further research.

2.1 Introduction

This chapter reviews the data collection systems for child mortality data, the common measures of childhood mortality, the trends of child mortality in some African countries, the conventional methods of measuring child mortality, the effect of HIV on these methods, and methods that have been developed to estimate and correct the bias due to HIV on child mortality estimates. The conclusion of this chapter gives an insight into the vacuum that still exists in the methods of estimating child mortality and explains the part to which this research aims to contribute.

2.2 Child mortality data collection systems

A number of approaches can be used to collect data for calculation of infant and child mortality rates.

2.2.1 Vital Registration

Complete vital registration systems are by far the best source of accurate demographic data that can be used to assess child mortality levels and trends. However, in most sub-Saharan countries the vital registration systems are unusable due to lack of completeness. Even in countries whose registration systems are of reasonable completeness, the data for computation of child mortality delays in registration may affect the timeliness of the estimates (Hill, 1991). Delayed registration may also lead to underreporting of child mortality since in most cases deaths of young children are seldom registered if births of such children were not initially registered.

2.2.2 Surveys

The inability of vital registration systems to provide adequate statistics for the computation of demographic measures has left surveys that collect birth history data as the primary source of data for the calculation of infant and child mortality. A number of surveys have been used to collect infant and child mortality data over the years; these include Multiple Indicator Cluster surveys (MICS), World Fertility Survey (WFS) and the Demographic and Health Surveys (DHS). Full birth history and summary birth history data are the main forms in which data is extracted from surveys. Full birth histories capture, for each woman of reproductive age, comprehensive information for each live birth and death (if any) of each child. Women are asked about the date of birth of every

live-born child she has ever had; she is also asked a question on whether each child born is still alive and if not, she has to provide the date of death of the child. Complete birth histories have been widely used in developing countries for both the Demographic and Health Survey (DHS) and its predecessor, the World Fertility Survey (WFS). The data obtained from a full birth history can be used to compute life tables for children for five-year or ten-year periods prior to the survey date. The quality of estimates for more distant periods before the surveys can be compromised by event omission and misreporting of dates of births and deaths (Korenromp *et al.*, 2004; Hill, 1991; Bicego and Ahmad, 1996). Use of life tables helps to solve the problem of censoring which results from children who are not exposed to the risk of death for the full period under consideration.

There are various challenges associated with the collection of full birth history data. Besides the high costs involved in the administration of surveys that collect full birth history data, the complexity of the data also requires specially trained interviewers. Full birth history surveys have many questions in the child health section; the DHS, for instance, asks women eleven questions of each surviving child and nine questions for each dead child in the full birth history section alone (Rajaratnam *et al.*, 2010b). Surveys that collect full birth history data normally have small sample sizes due to their high administration costs and their complicated nature; furthermore, they are designed to produce nationally representative estimates. This means that they fail to provide infant and child mortality information at the sub-national level and for small population groups.

Unlike full birth histories, summary birth histories involve only two questions for extracting data for the calculation of child mortality estimates. They require each woman of reproductive age to provide the number of children she has ever borne as well as the number of those children who are still surviving. The short and simple nature of the questions means that the data obtained is less prone to error compared to full birth history data, which are prone to recall errors, since few women can recall the exact dates of births and deaths of their children (Hill, 1991). There is also a tendency by interviewers to misdate events to avoid further questions, especially in surveys which collect full birth history data, like the DHS, where the survey instrument is long. Summary birth history questions can easily be incorporated in large exercises like the census resulting in the estimation of infant and child mortality estimates for small population groups and sub-populations (Rajaratnam *et al.*, 2010b).

2.3 Measurement of child mortality

Traditionally childhood mortality used to be assessed using a single measure, the infant mortality rate (IMR) (Hill, 1991; United Nations, 1990). Lack of complete vital registration systems and the realisation that up to fifty percent of child deaths in most low-income countries occur after age one, has led to the broadening of the age range considered for the computation of child mortality (Hill, 1991; United Nations, 1990). The probability of dying between birth and exact age five, $q(5)$, commonly referred to as the under-five mortality rate is now the preferred measure of child mortality. The under-five mortality rate is preferred, especially when using indirect demographic techniques, because it is more robust and is less sensitive to assumptions compared to the infant mortality rate (Ahmad *et al.*, 2000).

2.4 Child mortality trends

In many developing countries, data for producing estimates of mortality rates from which trends from the distant past can be derived are rarely available. Although availability of data has improved due to an increase in the number of surveys that collect data used for computation of childhood mortality estimates, there is still a challenge in producing trends using data from the different sources. The differences are mainly a result of the differences in data quality of the surveys and the methods and datasets used by different institutions to produce estimates, which result in sampling and non-sampling error (Ahmad *et al.*, 2000; Rajaratnam *et al.*, 2010a).

The under-five mortality trends for the countries used in this study have been obtained from various sources including: Demographic and Health Surveys, reports from the UN Inter-Agency Group for Child Mortality Estimation (IGME) website, which is a collaboration of UNICEF, WHO, the World Bank, UN Population Division and analyses by Ahmad *et al.* (2000).

The IGME derives childhood mortality estimates by applying standard methods to compute childhood mortality estimates and the time points at which they apply (Inter-agency Group for Child Mortality Estimation (IGME), 2010b). The choice of the method depends primarily on the nature of the data being used, for full birth history data, the IGME uses direct estimation procedures and for summary birth history data they use indirect estimation techniques. A smoothed series of childhood mortality estimates is then derived using estimates from all the sources and applying a regression

model. The choice of the regression model to be applied for a given country is based on the HIV prevalence level.

For low HIV prevalence countries the IGME uses a linear spline regression model. In this model the log of the U5MR or the IMR is regressed on time. Weights are assigned to the individual observations based on the accuracy of a particular type of observation according to a defined scale of weights. The weight values are determined based on data quality and consistency. The regression model used is:

$\ln({}_nq_0)_i = b_0 + b_1(date)_i + b_2(postk1)_i + b_3(postk2)_i + \dots + e_i$, where $\ln({}_nq_0)_i$ is the logarithm of either the U5MR or the IMR, *date* is the calendar year of each observation, *postk1* is *date* minus the date of the earliest defined knot if positive, or zero otherwise, *postk2* is *date* minus the date of the second earliest defined knot, b_k , ($k = 0, 1, 2, \dots$) and are the regression coefficients, which represent the rate of change of either the IMR or the U5MR with time in a particular period. The error term e_i ($i = 1, 2, 3, \dots$) is assumed to be normally distributed around the logarithm of the mortality indicator with mean zero and constant variance, σ^2 .

The IGME also developed a method for adjusting for bias in the empirical estimates due to HIV/AIDS for countries with high prevalence of HIV. A country is deemed to be in this category if it has an HIV prevalence level that has exceeded 5 percent in women aged 15-49 years. Currently, the available adjustment procedure is applicable only to estimates derived from full birth histories and estimates from summary birth histories are excluded from the estimation process.

The regression model for countries with high prevalence of HIV is implemented by first subtracting UNAIDS-WHO estimates of HIV/AIDS under-five mortality from all observations in the epidemic period. A regression model is fitted to the data points for under-five mortality from all causes other than HIV/AIDS. The series is extrapolated to a recent date of interest and to get the new series which include HIV/AIDS, the UNAIDS-WHO under-five mortality rates are added to the regression estimates in the epidemic period.

The estimates compiled by the IGME are useful; especially for assessing trends, however, the accuracy of the estimates depends on the reliability of the method used to compute the original estimates. In addition, the reliability of the estimates derived for high-HIV prevalence countries depends on the validity of the assumptions on the distribution of HIV-negative and HIV-positive birth to HIV-positive mothers which will depend on the stage of the epidemic and the magnitude of intervention in a country.

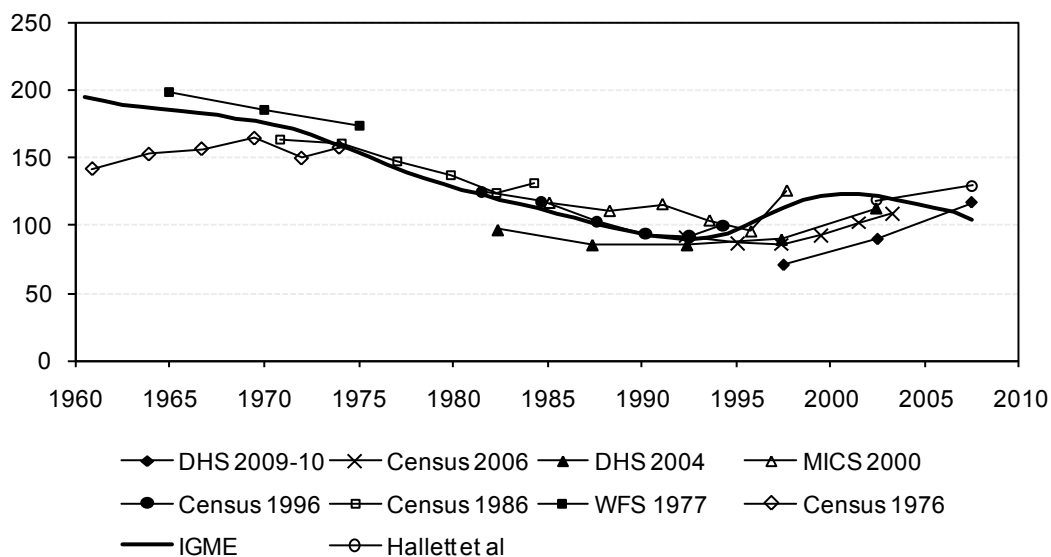
2.4.1 Trends in Under-Five Mortality Rate in Selected Countries

The following sub-sections present trends in under-five mortality rate in some sub-Saharan African countries that have been selected to study bias. The estimates presented have not been corrected for bias due to HIV. The general trend shown for the countries presented in Figures 2.1 - 2.6 indicate that generally under-five mortality declined from around 1960 to the late 1990s. However, in the years that followed, the rate of the decline became slower or even stalled (Adetunji and Bos, 2006). As from just before 1995, most countries show an increasing trend in under-five mortality. The period of under-five mortality either stalling or increasing could be explained by the impact of the HIV epidemic. The consistency of under-five mortality estimates is illustrated by how close to or dispersed from each other the estimates are. Estimates that are tightly grouped imply that the estimates are consistent, while widely scattered estimates suggest inconsistency or uncertainty in the under-five estimates.

2.4.1.1 *Under-five mortality estimates from various sources: Lesotho*

Figure 2.1 illustrates the under-five mortality levels and trends from different sources in Lesotho. The data show an overall decline in under-five mortality from 1960 to 2008. The under-five mortality rates in Lesotho decreased steadily in the 1970s through to the late 1980s, but the graph also shows some evidence of a rise in the late 1990s, probably indicating the effect of HIV. The estimates shown do not indicate great consistency, which might be explained by the use of different methods in deriving the estimates. The advent of AIDS resulted in excess risk of dying in infected children, which could explain the increase in under-five deaths from around 1995. However, the IGME projections predict a decline in the U5MR; they probably anticipate increased efforts in prevention of mother-to-child transmission of HIV and an increase in the rollout of antiretroviral drugs.

Figure 2.1 Under-5 mortality estimates from various sources: Lesotho

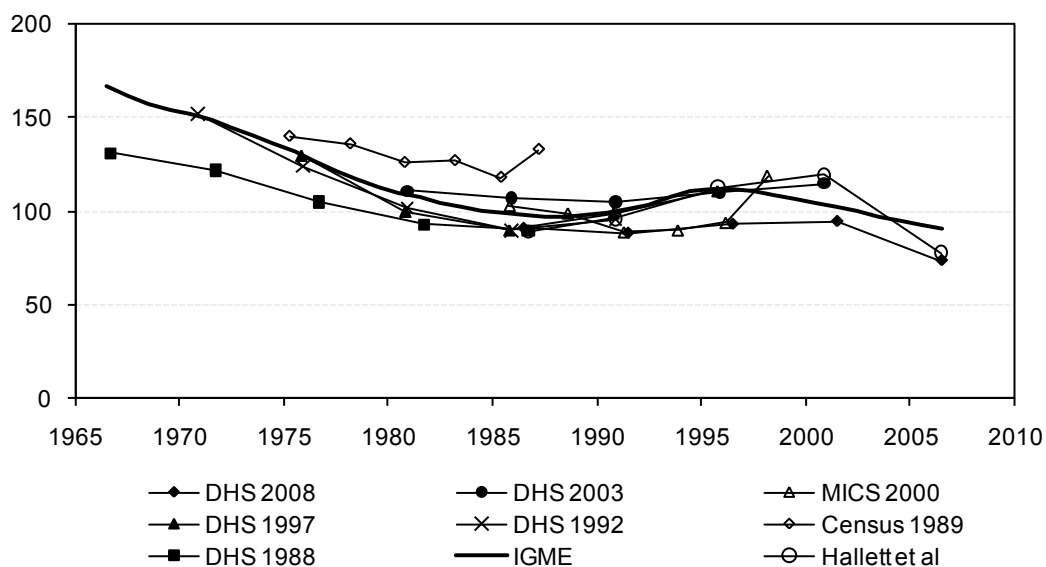


Source: IGME (2010a); Hallett *et al.* (2010)

2.4.1.2 Under-five mortality estimates from various sources: Kenya

The estimates of under-five mortality for Kenya from all the sources shown suggest an overall decline from around 1965 to 2005. Although Kenya experienced a surge in the U5MR in the 1990s, as shown in Figure 2.2, recent surveys suggest a decline in the U5MR. The recent downward trend could be explained by the efforts that are being made to mitigate the effects of HIV or could also be a result of the bias, which is introduced by HIV.

Figure 2.2 Under-5 mortality estimates from various sources: Kenya

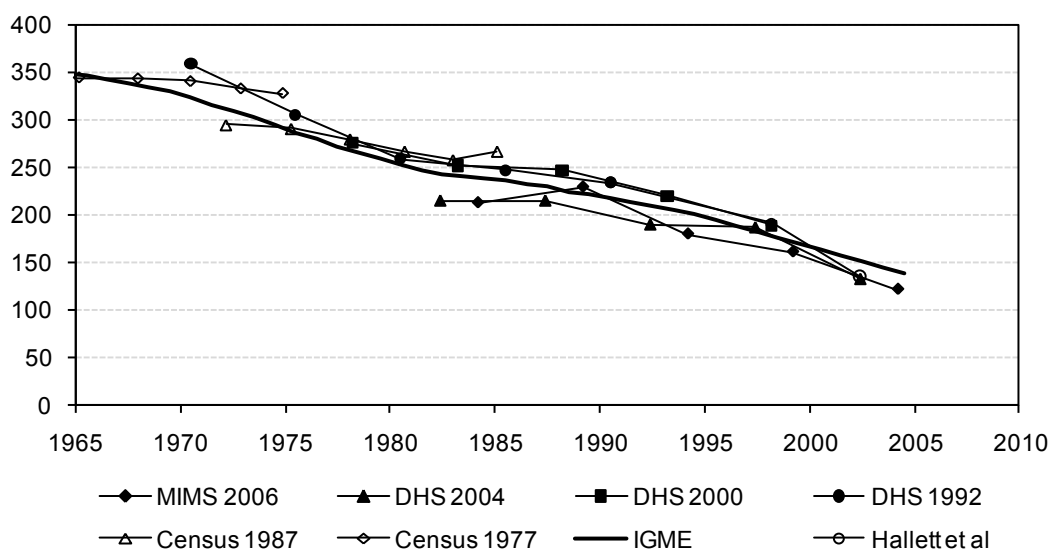


Source: IGME (2010a); Hallett *et al.* (2010)

2.4.1.3 Under-five mortality estimates from various sources: Malawi

Figure 2.3 shows that Malawi experienced a consistent decline in under-five mortality rates from around 350 deaths per 1000 births to just over 150 deaths per 1000 births around 2005. Unlike the data from the other countries, the Malawian data do not suggest a clear picture of the suspected HIV effect, which is generally shown by a slight rise in infant and under-five mortality rates around the late 1990s. This is likely because Malawi has much higher background under-five mortality, resulting in a small proportion of the childhood deaths being attributed to HIV.

Figure 2.3 Under-5 mortality estimates from various sources: Malawi

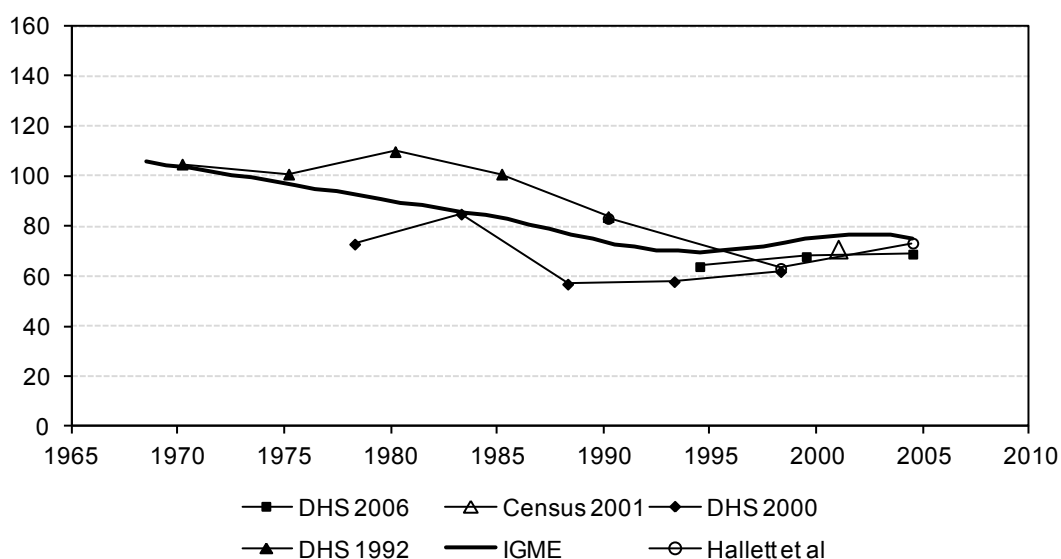


Source: IGME (2010a); Hallett *et al.* (2010)

2.4.1.4 Under-five mortality estimates from various sources: Namibia

The under-five mortality rates from Namibia shown in Figure 2.4 are much lower than the other countries in this study. However, the pattern shown in Figure 2.4 shows minimal decline in the infant mortality rate and the most recent estimates derived from the 2006 DHS survey even show an increase in the under-five mortality rate. The under-five mortality rate increased from 38 deaths per 1000 births in the 0-4 year period before the 2000 DHS to 46 births per 1000 deaths in the 0-4 year period before the 2006 DHS. Unlike in other countries, the estimates for Namibia from the different sources are not very close to each other, suggesting lack of consistency among estimates possibly as a result of the impact of HIV.

Figure 2.4 Under-5 mortality estimates from various sources: Namibia

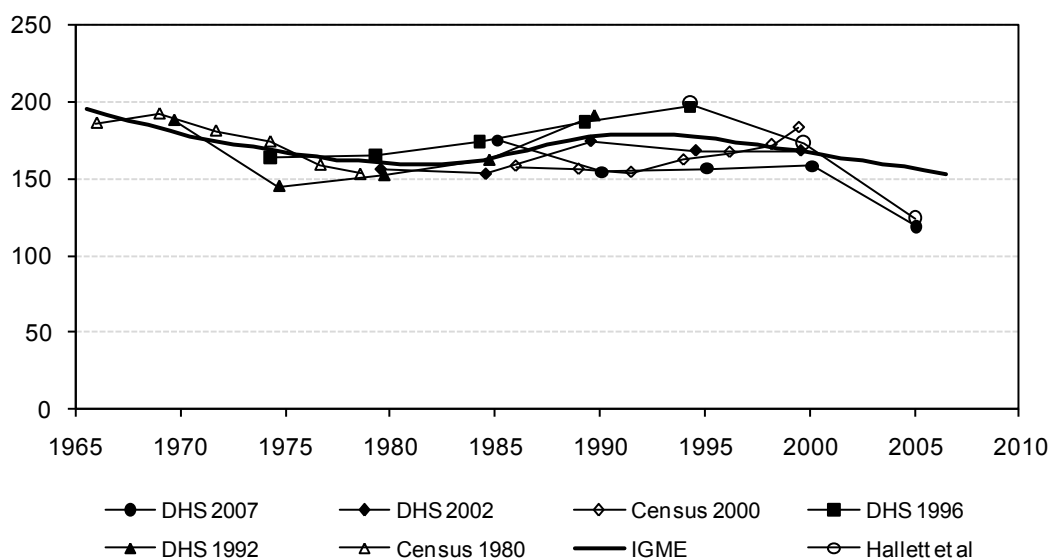


Source: IGME (2010a); Hallett *et al.* (2010)

2.4.1.5 Under-five mortality estimates from various sources: Zambia

Figure 2.5 shows estimates of under-five mortality from various sources for Zambia. The estimates show an overall decline in the level of under-five mortality from just under 200 deaths per 1000 births to around 120 deaths per 1000 births from 1965 to around 2005. However, the decline has not been consistent over the time from around 1965 to around 2005. In the period from around 1980 to 2000, under-five mortality remained constant in the earlier years leading to 1995 at around 150 deaths per 1000 births, thereafter, the mortality level rose probably in response to the HIV epidemic. Recent estimates show signs of a slight decline in the level of under-five mortality in Zambia. This is probably a result of the response to the HIV epidemic through treatment and prevention efforts.

Figure 2.5 Under-5 mortality estimates from various sources: Zambia

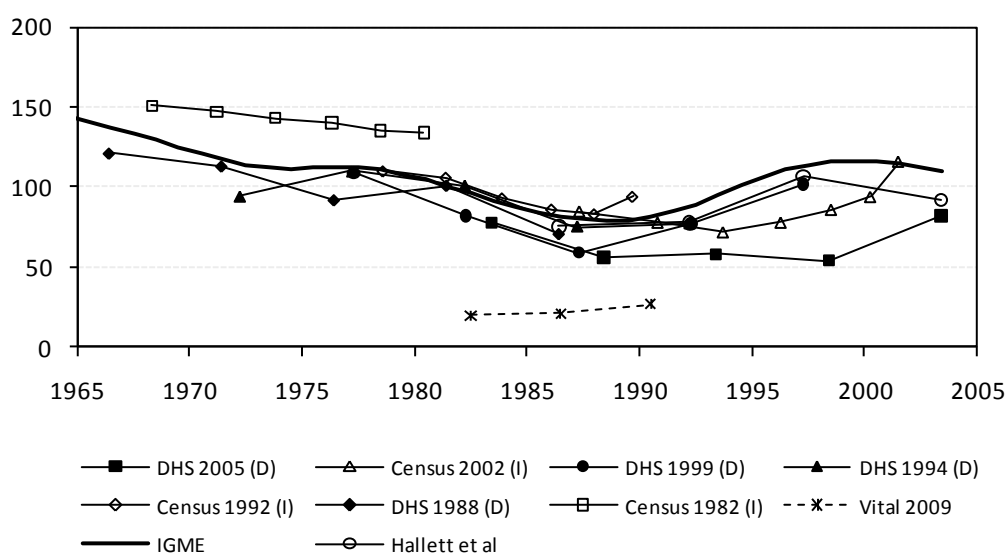


Source: IGME (2010a); Hallett *et al.* (2010)

2.4.1.6 Under-five mortality estimates from various sources: Zimbabwe

The under-five mortality estimates from Zimbabwe show that there has been a steady decline over the years from 1965 to the late 1980s. The trend reverses from the late 1990s with evidence of an increase in the level of under-five mortality as shown in Figure 2.6; this again could be explained by HIV/AIDS epidemic. The estimates from vital registration are very low compared with estimates from other sources probably because of the under-registration of child deaths that is synonymous with this data source.

Figure 2.6 Under-5 mortality estimates from various sources: Zimbabwe



Source: IGME (2010a); Hallett *et al.* (2010)

2.5 Approaches to estimating child mortality

Infant and child mortality can be estimated by using the direct estimation approach or by applying indirect methods. The choice of an approach to use in estimating infant and child mortality largely depends on the available data and the available resources for the collection of data. Direct estimation of child mortality utilises full birth history data obtained from specifically designed surveys like the DHS or from vital registration systems whereas indirect methods are commonly applied to data obtained from censuses and other general surveys (Rutstein and Rojas, 2003). The following sections outline these methods and conditions under which the methods are applied.

2.5.1 Direct estimation of child mortality

To estimate mortality rates directly from data on full birth histories, the life table approach is used. The life table method calculates probabilities of dying based on the reported dates of birth and death and the numbers of children of a particular age exposed to the risk of dying during a specified period (Bicego and Ahmad, 1996). The approach uses synthetic cohorts, where the total exposure to the risk of dying consists of contributions to person-years of risk of children of different birth cohorts. Unlike the real cohort life table approach, which has limitations on calculating full exposure especially for under-five mortality rates since only data on children born five or more years before the survey can be utilised, the use of synthetic cohorts allows for the use of recent data (Rutstein and Rojas, 2003).

2.5.2 Indirect methods for estimating child mortality

Many developing countries, as already highlighted in preceding sections, lack complete or accurate data for computing descriptive measures of demographic processes. Ideally, data for vital events can be obtained from civil vital registration systems and data obtained from general surveys like censuses could be used to calculate the exposure to risk. Though censuses and other surveys provide alternative sources for calculating demographic measures in populations with deficient vital registration systems, the data obtainable from such sources is not directly usable due to coverage and content errors like under-enumeration and reporting errors.

2.5.2.1 *The Children ever born/Children surviving (Brass) method*

Brass developed a widely-used indirect technique for calculating child mortality from reports by women of reproductive age on the number of their children born alive and the number that are dead (Brass *et al.*, 1968). These retrospective reports can be

extracted from many surveys and they are widely thought to be accurate compared to data which require women to recall the exact dates of the births and/or deaths of their children. Brass developed the method based on the relationship between the proportion of children dead and the overall level of childhood mortality. The method permits the calculation of $q(x)$, the probability of dying between birth and exact age x , by allowing for the duration of exposure to the risk of dying (United Nations, 1990). Although the method has been refined a number of times since its first formulation, it still remains the best choice for demographers in calculating the childhood mortality estimates for developing countries (Trussell, 1975; Sullivan, 1972; Feeney, 1980).

The method converts the proportions of dead children among children ever born by women in reproductive age groups, that is, 15-19, 20-24, . . . , 45-49, into probability estimates of dying before reaching exact childhood ages. The proportions dead among children ever born to women aged i is given by:

$$d_i = \frac{D_i}{B_i} = \frac{\int_0^{\alpha} D_i(a) da}{\int_0^{\alpha} B_i(a) da}, \text{ where} \quad (1)$$

B_i is the total number of children born to women aged i at the time of survey, D_i is the total number dead, among children born to women aged i at time of survey, $D_i(a)$ is the number of deaths among children born a years before the survey to women aged i at the time of the survey, $B_i(a)$ is the number of births a years before the survey to woman aged i at time of survey, and α is the earliest age of childbearing.

The relationship between the level of childhood mortality and the proportion of children dead among children ever born is not exact (Preston *et al.*, 2001). There are other factors such as the distribution of reproductive histories for the women who are reporting their birth, which affect childhood mortality. To compensate for the effect of non-mortality factors on the overall probability of dying between birth and exact age x and the effect of the duration of exposure to the risk of dying, Brass developed a set of adjustment factors, $k(i)$. These adjustment factors were developed based on combinations of simulated fertility and child mortality schedules (United Nations, 1983; United Nations, 1990; Preston *et al.*, 2001). The adjustment factors can be computed using the following equation;

$$k(i) = a(i) + b(i) \frac{P(1)}{P(2)} + c(i) \frac{P(2)}{P(3)}, \text{ where} \quad (2)$$

$a(i)$, $b(i)$ and $c(i)$ are coefficients obtained from the regression of the mortality and fertility estimates generated from projected stable populations, $k(i)$ is the adjustment factor corresponding to women in the age group i and $P(i)$ is the average parity reported by women in age group i . The coefficients are derived for each family of model life tables which means that the mortality estimates derived using the CEB/CS method are also affected by the choice of model life table at this stage too.

The proportion dead among children ever born can be converted to life table probabilities of dying between birth and exact age x using the following equation:

$$q(x) = k(i)d(i), \text{ where} \quad (3)$$

$q(x)$ is the probability of dying between birth and exact age x , $d(i)$ is the proportion dead among children ever born to women in childbearing age group i ($i = 1, 2, 3, \dots, 7$) and $k(i)$ is the multiplier meant to adjust for non-mortality factors determining the value of $d(i)$.

The implementation of the Brass method is based on several assumptions that may not always hold and this has implications on the accuracy of the estimates derived. One assumption is that the mortality of a child depends on the age of the child only and not on other factors like socio-economic background, birth order or the age of the mother. This assumption does not always hold given that children born to younger women tend to experience an excess risk of dying due to premature birth and a lower birth weight than those children born to older women, which means that the risk of childhood death can also depend on mother's age. The estimates derived from information from women in the 15-19 year age group are likely to be upwardly biased and lead to unreliable recent estimates, hence they are normally disregarded (United Nations, 1983).

The Brass method also assumes that age-specific fertility and mortality rates have remained constant during the 30 or 35 years before the census or survey (Brass *et al.*, 1968). Changes in fertility will affect the parity ratios; if fertility is declining the mortality estimates will be upwardly biased because the exposure to risk of dying of children will appear to be more recent. The effects of fertility change, however, do not have any significant effect especially in Africa where changes in fertility are gradual hence they are usually ignored (Brass *et al.*, 1968). Also, the multipliers used in the CEB/CS method are not sensitive to small differences in average parities.

The adjustment factors and the reference time to which the estimates of mortality apply are obtained by making use of model life tables. The application of the

Brass method is based on the assumption that the age pattern of mortality from a given population resembles that of the chosen model life table. These model life tables were developed based on mortality patterns from particular countries with unique mortality experiences thus they may not be appropriate for use in modern day sub-Saharan countries because of several factors including HIV (Coale and Demeny, 1983; Preston *et al.*, 2001).

The assumption of constant mortality was relaxed through the refinement of the original Brass method by several authors (Coale and Trussell, 1974; Feeney, 1980). Feeney (1980) developed an estimation procedure to establish the set of years to which infant mortality estimated from summary birth history data refer. Coale and Trussell used the following equation for the estimation of the reference time period to which $q(x)$ refers:

$$t(i) = a(i) + b(i) \frac{P(1)}{P(2)} + c(i) \frac{P(2)}{P(3)}, \text{ where} \quad (4)$$

$a(i)$, $b(i)$ and $c(i)$ are coefficients for each of Coale-Demeny model life tables, obtained from regression on the average parities obtained from simulated fertility schedules, $t(i)$ is the reference time for each estimate derived from women of age group i , which can be converted to the reference date by subtraction from the date of the survey or census, and $P(i)$ is the average parity reported by women in age group i .

Feeney (1980) used simulation to develop an estimation procedure that relaxes the assumption that mortality was constant over time and allows for the calculation of time to which the estimates refer. The variant by Feeney is, however, not preferred because it was developed based on $q(1)$ which is sensitive to the underlying age pattern of mortality and can produce biased estimates if the childhood mortality pattern of the population differs from the pattern in the general standard (United Nations, 1983).

The other assumption, which is perhaps the most important for this study, is that the survival of children is independent of that of their mothers. This assumption is violated in populations with a generalised HIV epidemic, more so if the rates of breastfeeding are high. In these populations, HIV is passed from the mothers to their children through vertical transmission. Estimates put the chance of transmission of HIV from the mother to her child between 25-45 percent. This HIV effect has threatened the applicability of the Brass method because it negatively affects the method in numerous ways. So far, there has not been an acceptable solution to the threats posed to the Brass method by HIV. The most significant attempt in this regard by Ward and Zaba (2008) is

of limited use mainly because of the assumption of a stable population and stable HIV prevalence, since the epidemic can evolve through time and the age pattern of mortality also changes. This method is dealt with in detail in section 2.8.1.

Despite the problems that can be faced in the implementation of the Brass method, it will remain the primary source of information about child mortality, although information on household deaths may prove usable, as may sibling histories. The simple nature of the questions means that the chances of response errors relative to full birth history questions are low and it means that they can be included in most surveys. This means that indirect methods used to compute estimates from summary birth history data and direct estimation techniques used to derive mortality estimates from full birth history data obtained from household surveys remain the primary source of childhood mortality estimates, so there is the need to continuously refine the methods so that they yield reliable estimates.

2.6 Relative bias in methods of estimating childhood mortality

In the absence of HIV and data problems, infant and under-five mortality estimates derived from both the indirect method and directly from full birth history data are expected to approximate the true value. Studies have, however, shown that this seldom happens as the infant and child mortality estimates derived using the two approaches often differ. In most cases, especially in child estimates derived from younger women, infant and under-five mortality estimates derived from indirect methods are higher than estimates calculated directly from full birth history data (Adetunji, 1996; Hill, 1991). Adetunji (1996) used data from selected African countries to show that infant mortality estimates derived from the indirect method were much higher than estimates obtained directly from full birth history data for the same age groups of women. A combination of data errors, violation of assumptions and intrinsic biases in the method were found to be the factors that explain the difference.

However, there has been some debate as to which method gives the best estimates of child mortality. Some researchers have favoured the use of the child estimates derived directly from full birth histories because of the fact that they depend on fewer assumptions, while others have preferred the indirect methods for the robust estimates they produce in populations with defective data (Knodel and Chamrathirong, 1978; Hill, 1991; United Nations, 1992).

The intrinsic differences between the methods make it difficult to analyse trends in child mortality levels since available estimates will have been derived from different

methods. Comparing estimates derived from these two methods or pooling their estimates for developing trends without any form of adjustment will lead to wrong conclusions since the true level of child mortality is difficult to determine.

2.7 Effect of HIV on the methods of estimating child mortality

HIV-positive women are at a higher risk of dying than women who are uninfected (Mahy, 2003). Children born to HIV-positive mothers have an excess risk of dying due to AIDS as a result of vertical transmission or because of an indirect increase in mortality if the mother is sick or has died (Mahy, 2003; Ng'weshemi *et al.*, 2003; Marston *et al.*, 2005). The proportion of children of HIV-positive mothers who become infected in the absence of anti-retroviral therapy or PMTCT is between 25 and 45 percent (Marston *et al.*, 2005; Timaeus, 1998; Zaba *et al.*, 2003). The proportion, however, reduces to around 20 percent in populations where there is no breastfeeding (Adetunji, 2000).

Infant and child mortality data for children born to HIV-positive mothers is likely to be missed by surveys since HIV-positive mothers are at a high risk of dying between the time of giving birth and the survey date. This results in infant and child mortality estimates that are downwardly biased since data for these children are likely not to be included in the estimation of infant and child mortality. Omission of data on children born to HIV-positive women affects both direct and indirect estimation of infant and child mortality in the same way (Rutstein and Rojas, 2003). The bias due to HIV in infant and child mortality estimates might be expected to be lower for recent estimates since it increases with time since the mothers' infection (Artzrouni and Zaba, 2003).

Indirect techniques suffer more from the effect of HIV as compared to direct estimation procedures partly because of the assumptions that are necessary for the implementation of the technique. The CEB/CS technique assumes that the mortality of children is independent of that of their mothers. In HIV populations, this assumption is violated because of vertical transmission of HIV from the mother to the child as mentioned earlier in this section.

In contrast to direct estimation techniques, indirect techniques use model life tables for the conversion from the proportion dead children to the probability of dying by age x . In particular, the adjustment factors used in the conversion of the proportion dead among children ever born to the actual life table probability of dying between birth and exact age x were developed from simulations of particular families of model life

tables (United Nations, 1990). These life tables were developed from observed mortality risks in HIV-free populations and since HIV alters the age-specific mortality patterns in children, the use of these life tables in HIV populations lead to biased estimates.

HIV introduces heterogeneity in child mortality with respect to the age of the mother thus violating the assumption that the mortality of children of women in particular age groups represent mortality of the children of women in the whole population since HIV prevalence varies with age (Zaba *et al.*, 2003; Mahy, 2003). This also violates the assumption that the mortality of the children depend on their ages only.

The estimates of the time at which the rates apply are also affected since they also use model life tables, which are unrepresentative of mortality patterns, observed in populations with generalised epidemics of HIV. Even for direct estimation, Artzrouni and Zaba (2003) observed that the bias increases as estimates refer further back in the past and hence they recommended that estimates based on birth histories should be calculated for recent periods.

The conversion of estimates of $q(x)$ values to a common index, for example to $q(1)$ or $q(5)$, is also based on the use of a model life table whose age pattern of mortality may not resemble that prevailing in a population with a generalised HIV epidemic.

The error introduced into the estimation methods by HIV means that if reliable estimates are to be obtained from the data there is a need for refinement of the methods through the development of appropriate adjustment factors. The impact of HIV on indirect methods of estimation seems to be greater than that on direct estimation. This is unfortunate since data for indirect methods is easier to obtain than full birth history data used for direct estimation.

2.8 Methods for correcting for HIV bias on child mortality estimates

Part of the focus of recent research on the impact of HIV on child mortality has been on refining methods of measuring child mortality to account for the bias that results from the impact HIV. These efforts have been buoyed by the realisation of the threats posed by HIV on measures of child mortality. In the next sub-section, the methods that have been developed to estimate bias introduced by HIV on child mortality estimates are reviewed. In particular, the review is centred on the research done by Ward and Zaba (2008), the recent work of Hill and Walker and Hallett *et al.* (2010).

2.8.1 Ward and Zaba correction

Ward and Zaba (2003) assessed the extent of bias introduced by HIV on child mortality estimates derived from the CEB/CS method by exposing simulated stable populations to various levels of HIV prevalence in women of reproductive age. They calculated the proportion dead from the projected populations whose mortality, fertility and HIV incidence schedules are known. These proportions of dead children are used to calculate the level of child mortality for each simulated population using the CEB/CS method.

The magnitude of the error introduced by HIV into the CEB/CS method is obtained by comparing the estimated child mortality rates, $q(z)^e$ derived from the CEB/CS method with the true child mortality rates $q(z)^t$ as follows:

$$n(z) = q(z)^t - q(z)^e, \quad (5)$$

where $n(z)$ is the size of the error and the other components in (5) are as defined above.

Apart from measuring the extent of bias introduced by HIV on the CEB/CS method, Ward and Zaba (2008) developed sets of correction factors that can be applied to child mortality estimates obtained from HIV populations to compensate for the effect of HIV. The factors for correcting estimates are derived from regression modelling of the bias against HIV prevalence.

The problem with the Ward and Zaba method is that they assumed stable populations and stable incidence of the HIV epidemic over time. Observations made in most countries affected by the HIV epidemic reveal that the incidence levels change depending on the stage of the epidemic and the extent of interventions in terms of anti-retroviral therapy provision and other prevention strategies (Mahy, 2003; Garenne and Gakusi, 2006).

2.8.2 Hallett *et al.* correction

Hallett *et al.* (2010) used child mortality and fertility history data including births to women who subsequently died, collected through verbal autopsy, from longitudinal household surveys in eastern Zimbabwe during the period from 1998 to 2005 to estimate the extent of bias introduced by HIV. They used the results from the longitudinal survey to develop a mathematical model that estimates the extent of bias due to the correlation between the mortality of the mothers and their children in countries affected by the HIV epidemic using data obtained from national surveys. The model can also be used to derive corrected infant and under-five mortality rates and

their trends in populations affected by the HIV epidemic. The model was then re-parameterised and used to estimate and compare bias of estimates obtained from demographic and health survey (DHS) data from six other sub-Saharan African countries to determine the behaviour of the bias in environments with different background infant and child mortality and different levels of HIV-related epidemic profiles.

For the longitudinal household study, Hallett *et al.* (2010) used standard DHS questions and procedures. To obtain child mortality and fertility history data of deceased women, verbal autopsy interviews were conducted of the caregivers of the women who died since the last interview. Such data included births that occurred between the date of the last round and the date of the deaths of the cohort member and information on any children who had died both among those born since the last round or before it.

The data on child mortality and fertility history of deceased women enabled the derivation of infant and child mortality estimates that are corrected for under-reporting because of HIV-related deaths to women. These corrected estimates were calculated by combining the estimates from children born to both surviving and deceased women with adjustments being made for the sampling of women in the open cohort and for censoring. Data for deceased women were adjusted for censoring with the underlying assumption that deceased women would have been lost to follow-up at the same rate as surviving women for reasons other than death. This assumption is likely to be violated since women who survive are likely to have a higher risk of moving to seek employment, for example, so this may lead to lower estimates of bias. Hallett *et al.* (2010) showed that violation of this assumption will lead to changes in the estimates of bias. They showed that if the loss to follow-up for women who died is greater than that for women who survived then the bias would increase.

The model is an individual-based stochastic model that simulates the mortality and fertility experiences of women born in rural Zimbabwe between 1920 and 2005. The data used to set the assumptions of the model were drawn from several studies including from countries outside Zimbabwe. For fertility and mortality assumptions, data from surveys carried out in rural Zimbabwe in the pre-AIDS era were used. The size of the births cohorts were estimated from United Nations Population Division data. The model also allowed for the use of anti-retroviral therapy for women with CD4 count less than 200 per micro litre of peripheral blood, with mean time to death on treatment initiation equal to ten years. Probabilities of mother-to-child-transmission

were computed based on the data from studies done in Zimbabwe. Information for HIV sub-fertility assumption was obtained from studies from rural Uganda.

The extent of the bias or under-estimation of infant and child mortality is taken to be the difference between the corrected rates, which include child mortality and birth history data of deceased women, and the uncorrected rates, which are obtained in the standard manner. Analyses of the extent of the biases over time from the countries where the model was applied show that the bias varies according to the size and stage of the epidemic of HIV infection and background infant and child mortality with greater bias being observed in populations where the epidemic is in later stages and background mortality is low.

Although the model developed by Hallett *et al.* (2010) has been used to estimate the extent of bias introduced by HIV on estimates derived from full birth history data, it has not been used to assess the extent of bias introduced by HIV to child mortality estimates derived using the widely used CEB/CS method. Indirect techniques, regardless of their shortcomings might still remain the preferred choice among methods of estimating demographic measures given that in most developing countries very little progress has been made in improving the vital registration systems. Furthermore, summary birth history data will be available from the current round of censuses and if adjustments are not made to allow for the impact of HIV, the Central Statistics Offices will likely apply the method as it is. This means that efforts should be made in modifying existing methods or even developing robust indirect estimation that produce as accurate results as possible.

2.8.3 The IGME (Hill and Walker) method

The IGME developed a method for adjusting for bias in childhood mortality estimates derived from full birth history data due to HIV/AIDS for countries with high-prevalence of HIV (Inter-agency Group for Child Mortality Estimation, 2010b; Inter-agency Group for Child Mortality Estimation, 2010a). Countries with an HIV prevalence level that exceeds 5 percent in women aged 15-49 years are adjusted for this bias. The correction procedure is aided by the use of an Excel-based workbook¹. This method is also referred to, in this study, as the Hill and Walker model.

The method uses estimates of annual births and HIV prevalence among pregnant women aged 15-49 years, obtained from the latest projection of a national

¹ The workbook for correcting under-five mortality estimates derived from full birth history data for HIV bias was provided by Professor K. Hill through personal communication.

population and its epidemic, to estimate bias in survey estimates of under-five mortality for periods 0-4, 5-9 and 10-14 years before the survey. The method involves finding estimates of reported births and under-five deaths, which are then compared with corresponding true births and deaths to produce the estimate of bias.

The following sections give an outline of the procedure for using the Hill and Walker model in estimating the extent of under-estimation in survey estimates of under-five mortality.

2.8.3.1 Estimating true births and true under-five deaths

To calculate true births and under-five deaths, the annual births are disaggregated into three components which are: births to HIV-negative women, which are all assumed to be HIV-negative; HIV-negative births to HIV-positive women; and HIV-positive births to HIV-positive women. Apportionment of births to HIV-positive women is based on the assumption that 35 per cent of the births are HIV-negative and that no transmission of the virus occurs through breastfeeding.

For each birth component, corresponding deaths for the following five years are estimated. The risk of dying of HIV-negative births, regardless of the HIV status of the mother, are obtained from the West family of the Coale-Demeny model life tables with an under-five mortality that best approximates the pre-AIDS under-five mortality in a population. A mortality risk schedule from cohort studies with a probability of dying by age 5 of 0.625 is used to estimate the deaths related to HIV-positive births. This is based on the assumption that the roll-out of ARV treatment is negligible.

2.8.3.2 Estimating reported births and under-five deaths

In this model, HIV-negative women are assumed to have zero probability of dying implying that the births and under-five deaths of children born to these women are all reported. Reported births and under-five deaths of children are calculated based on annual probabilities of surviving of HIV-positive women derived from a survival curve with a median survival time of 9.5 years on the assumption that mothers were, on average, infected for four years at the time of births of their children.

The true and the reported births and deaths are summed for each five-year period before a survey and the estimate of bias is computed by dividing the ratio of reported under-five deaths to reported births by the ratio of true under-five deaths to true births.

The Hill and Walker model also calculates estimates of bias to be used in cases where the full birth history data from a DHS survey shows evidence of birth

transference. In cases where birth transference is evident, the Hill and Walker model calculates appropriate estimates of bias which correct under-five mortality estimates for the periods between the calendar year of the survey and $t - 1$ years before the survey, $t - 2$ and $t - 6$ years before the survey and $t - 9$ and $t - 11$ years before the survey.

The survey estimates can then be divided by the bias for each five-year period before the survey to give corrected/adjusted under-five mortality rates.

University of Cape Town

3 DATA SOURCES AND METHODS

This chapter presents the data sources, data quality and the methods used to estimate infant and child mortality as well as the methods for estimating bias in infant and child mortality due to HIV/AIDS as used in this study.

3.1 Data Sources

Data used in this research come from Demographic and Health Surveys undertaken in the six countries considered in this study. The countries are Kenya, Lesotho, Malawi, Namibia, Zambia and Zimbabwe. The number of Demographic and Health Surveys carried out in each of the countries varied from two to five, with Lesotho having the least number of surveys while Kenya had the most surveys.

Demographic and Health Surveys are intended to be nationally representative, with samples usually ranging from 5 000 to 30 000 households. To make the samples nationally representative some districts may be over-sampled and weights are used to account for those over-sampled units. To enable comparisons of estimates between countries, standard protocols are followed for each aspect of the survey from data collection to data processing, including the use of standard questionnaires though individual countries may modify the questionnaires to include sections of their own interests.

In the Demographic and Health Surveys, a complete birth history approach is used to collect information that is used for direct estimation of infant and child mortality rates. In the complete birth history approach, women of reproductive age are asked the date of birth of each live birth they have had. They are also asked if each child is still alive, and if the child is dead they are asked to provide the date of death and age at death for that child. The DHS surveys used for the six countries in the study are listed in Appendix E.

3.2 Data Quality

The quality of infant and under-five mortality rates derived from retrospective data depends on the completeness of reporting of births and deaths. Misreporting of age at death affects direct estimates of child mortality by distorting the age pattern of mortality. The estimates of mortality will be biased, especially if the misreporting involves the transference from one age bracket to the other. Displacement of births also leads to incorrect direct estimates of infant and under-five mortality, especially if the

displacement is greater for dead children than living children. This problem will be assessed in sub-section 3.2.1.

The DHS datasets used were checked to see if comparable cohorts of women were interviewed in consecutive surveys by comparing the age distributions of women (see Appendix F). While there was reasonable consistency in age distributions of women aged 15 to 49 in consecutive DHSs from countries such as Zimbabwe and Kenya, the same conclusion could not be reached for Zambia, Malawi, Lesotho and Namibia.

The Malawian DHS for 1992 shows a different age distribution to the age distributions derived from the 2000 and 2004 DHSs. It is, however, not easy to compare the 1992 age distribution with the other DHSs because of the irregular time between the surveys. The 2000 and 2004 Malawian DHSs also show some inconsistency, for example, the size of the 20-24 year age group in the 2004 DHS is greater than the size of the 15-19 year age group in the 2000 DHS implying that there were sampling problems in the two DHSs.

The two DHSs from Lesotho show some sampling problems as indicated by differences in the birth cohorts of women sampled. The proportions of women in almost all birth cohorts are lower in the 2004 DHS than in the 2009 DHS. The Namibian DHS also show sampling problems since in most cases the proportion of women in the same cohort is lower in an earlier survey compared to the proportion shown in a more recent survey.

Estimates of infant and under-five mortality derived from the CEB/CS method are also affected by the reporting of the mother's age, since the data are classified by mothers' age. In this study, no adjustments have been made to the DHS data apart from the adjustments that were done by Measure DHS before the release of the data. The data used in this study were investigated for the problem of births transference and the indices of births transference are presented in the next-subsection.

3.2.1 **Birth Transference**

Birth transference is a systematic error which results when births of children born t years before the survey are deliberately recorded as having occurred earlier than their actual date. The DHS surveys involves lengthy questionnaires for children born after certain specified dates and interviewers keen to reduce their workload incorrectly record some births as having occurred before the cut-off date, when in fact, such births would have actually occurred after the specified date t . Birth transference is usually more pronounced in deceased children than in surviving children (because of desire to avoid

having to ask questions about children who have died) and it often results in downwardly biased estimates of mortality and can also increase mortality estimates in the earlier period as a result of a shift in the derived ages at death. The transference of births will also result in distortions in the derived ages at death. The mortality estimates for the period close to the survey derived using these incorrect ages will be lower since some deaths which were supposed to contribute to mortality estimates for this period will have been moved to the earlier period.

In surveys with more serious birth displacement among dead children, the mortality estimates calculated for the period 5-9 years before the survey is usually larger than the estimates calculated for the periods 0-4 and 10-14 years before the survey, all other things being equal.

The birth ratio used in DHS reports was used to check for the problem of birth transference in the DHS surveys that were used to calculate the corrected rates of infant and under-five mortality using full birth history data (Central Statistical Office [Zimbabwe] and Macro International Inc, 2007). The formula for calculating the ratio is:

$$\text{Birth ratio (BR)} = 2 \times B_t / (B_{t-1} + B_{t+1})$$

where, B_t is the number of births in the calendar year t of the cut-off date as reported in the DHS reports, B_{t-1} is the number of births in the calendar year before the cut-off date and B_{t+1} is the number of births in the calendar year after the cut-off date. The following criteria were used to judge the extent of birth transference in the DHS surveys:

- a. $BR \geq 0.95$, No birth displacement.
- b. $0.90 \leq BR < 0.95$, Little birth displacement problem.
- c. $BR < 0.90$, Serious birth displacement.

Table 3.1 presents the indices of births transference (presented in bold) that were calculated from the DHSs used to derive the corrected levels of infant and under-five mortality for each country under study. The mortality estimates are affected by the problem of birth transference, especially if the problem is more pronounced in dead children than in living children.

The calculated indices show that only the Zimbabwean DHS is not affected by birth transference since all the indices of births transference are greater than 0.95 regardless of the survival status of the children. Since the indices for the other countries show evidence of birth transference, their infant and under-five mortality estimates will be calculated for the 1-5 year period before the survey. This will minimize the effect of

births transference on mortality estimates since the transferred births will now be included in the mortality calculations.

Table 3.1 Births transference indices by country and DHS

Country	Cut-off Date	Year	Dead	Alive	Total
Kenya 2008 DHS	1 January 2003	2002	121	1226	1347
		2003	64	982	1045
		2004	83	1048	1131
		Index	0.62	0.86	0.84
Lesotho 2009 DHS	1 January 2004	2003	68	621	689
		2004	55	568	623
		2005	85	603	688
		Index	0.72	0.93	0.91
Malawi 2004 DHS	1 January 1999	1998	424	1718	2143
		1999	233	1342	1575
		2000	285	1816	2101
		Index	0.66	0.76	0.74
Namibia 2006-07 DHS	1 January 2001	2000	101	969	1070
		2001	53	797	850
		2002	62	812	874
		Index	0.66	0.89	0.87
Zambia 2007 DHS	1 January 2002	2001	144	927	1070
		2002	126	951	1077
		2003	144	1100	1245
		Index	0.87	0.94	0.93
Zimbabwe 2005-06 DHS	1 January 2000	1999	66	983	1048
		2000	81	928	1010
		2001	74	910	984
		Index	1.17	0.98	0.99

The estimates of infant and under-five mortality are calculated for the 1-5, 6-10 and 11-15 year periods before the survey instead of the 0-4, 5-9 and 10-14 year periods in order to minimise the effect of births transference for the affected countries.

3.3 Methods for estimating infant and child mortality

The methods used to estimate infant and under-five mortality using both full and summary birth history data are presented in the following sub-sections.

3.3.1 Estimation of infant and child mortality from full birth history data

DHS surveys provide full maternity history information, which is used to calculate direct estimates of infant and child mortality. The basic information required for calculation of direct estimates of mortality are: the date of birth and where applicable, the date of death of each child born to each woman of reproductive age. Mortality

estimates obtained from complete birth histories are generally thought to be of good quality (Hill, 1991; Curtis, 1995). Infant and under-five mortality estimates were calculated using a STATA version 11 program after adapting an SPSS code produced by Measure DHS downloaded from their website. The estimates produced by the Stata code were shown to be the same as those published in the reports; the code is presented in Appendix H.

The mortality rates from DHS full history data were calculated for each of the periods 0-4, 5-9, 10-14, 15-19 and 20-24 years before the survey using synthetic cohort probabilities of death. Each period consists of 60 months and is further divided into 8 age segments, that is, 0, 1-2, 3-5, 6-11, 12-23, 24-35, 36-47, and 48-59 months of completed age. The mortality rates are derived from the death probabilities calculated for each age segment. Each component of death probability is defined by an age interval and a time period.

Deaths occurring within a given age interval and time period are divided by the exposure of all the children who enter that subinterval alive to give the death probabilities for each age segment. Allocation of deaths to age segments is based on the assumptions that all deaths are assigned to the age group within which the children died if that age group is completed within one time period, and half the number of deaths are assigned to an age group if the age group the child died in begins in one time period and finishes in the next time period: the other half is thus allocated to the next age segment. An exception to this rule is made if the next time period in which an age group where a child died finishes later than the date of the survey; in this case all deaths are assigned to the death variable of the initial age group.

Exposure is contributed to any age segment in which a child enters alive. A rule similar to the one applied to the death variable, as mentioned in the previous paragraph, was applied to allocate exposure to age segments.

The mortality rates are then obtained using the survival probability of the children of all the relevant age groups.

3.4 Indirect estimation of child mortality

Indirect estimates of infant and under-five mortality were calculated using the Trussell variant of the Brass CEB/CS technique. The data for this procedure were extracted from the DHS data sets from the responses given by women in each of the reproductive age groups to questions on the number of children they have given live birth to and how many of those children are still alive or have died.

In this study one of either the North and West models of the Coale-Demeny model life tables is used, depending on which model life table is most appropriate for each country according to the available literature (Inter-agency Group for Child Mortality Estimation (IGME), 2010b; United Nations, 1983). The North model was generally chosen for most of the countries, that is, Zimbabwe, Zambia, Malawi and Kenya because it is appropriate for countries with high under-five mortality but relatively low infant mortality due to prolonged breastfeeding. Other organisations like the UNPD have also used the same model for calculation of indirect estimates of infant and child mortality for these countries (United Nations, 1983). The age patterns of mortality for Namibia and Lesotho were assumed to conform to the Coale-Demeny West life table because the other families of model life tables were not appropriate for the observed infant and mortality data.

3.5 Methods used to adjust infant and under-five mortality rates for underestimation due to HIV

The following subsections presents the methods used to derive the correction factors that were used to adjust the direct estimates derived from the DHS data for underestimation due to the impact of HIV on reporting of childhood mortality.

3.5.1 Hallett *et al.* method

Hallett *et al.* (2010) developed a model that uses HIV prevalence, estimates of fertility and estimates of mortality rates for the pre-AIDS period, to correct survey estimates for bias resulting mainly from the non-reporting of the mortality of children born to HIV-positive mothers who are likely to be dead or too ill at the time of the survey.

The fertility and mortality rates for uninfected individuals are obtained from DHS surveys conducted in a pre-AIDS era. In this research the baseline fertility and mortality rates are obtained from the first DHS survey in each country, that is, the 1989 DHS for Kenya, the 1988 DHS for Zimbabwe, 1992 DHS for Malawi, 1992 DHS for Zambia and for the 10-14 years period before the 2004 DHS for Lesotho since no DHS was conducted before the epidemic. The female birth cohorts from 1950 were obtained from the estimations published by the United Nations in the 2008 World Population Prospects (United Nations Population Division, 2009)

The HIV prevalence rates are obtained from the antenatal clinics surveillance surveys and DHSs for countries whose surveys involved HIV testing. That is, the 2004 DHS for Malawi, 2009-2010 DHS for Lesotho, 2003 DHS for Kenya, 2001-2002 and 2005-2006 DHSs for Zambia and Zimbabwe respectively.

3.5.2 Hill and Walker model

To implement the Hill and Walker adjustment methodology the reported births and deaths for children born to both HIV-negative and HIV-positive mothers are estimated.

The level of under-five mortality for HIV-negative births was estimated using the life tables which are chosen on the bases of the pre-AIDS mortality levels in a country. The pre-AIDS under-five mortality levels were estimated from the 0-4 year estimates from the following DHSs: 1989 DHS for Kenya, 1988 DHS for Zimbabwe, 1992 Malawi DHS, and 1992 Zambia DHS. For Lesotho the estimate 10-14 years period before the 2004 DHS was used since no DHS was conducted in the country before the epidemic.

3.5.2.1 *Adapting the Hill and Walker model to derive estimates of bias for infant mortality*

The model developed by Hill and Walker estimates bias in under-five mortality only. The model was adapted to estimate the bias in infant mortality derived from full birth history data by adjusting the true deaths and reported deaths. Since the workbook considered deaths from the year of births up to the fifth year in calculating the bias in under-five mortality estimates, the adapted workbook for infant mortality only considered deaths in the year of birth and in the first year after birth. The rest of the procedure was the same as the one for calculating the bias in under-five mortality outlined above.

3.6 Estimating bias in mortality estimates derived from the CEB/CS method

The overall bias in infant and under-five mortality estimates derived from the application of the standard CEB/CS technique comprises bias due to the impact of HIV on the survival of women and their children, the model life table used to convert the age-specific mortality rates to a index of mortality at a common age, bias due to timing and bias due to regression coefficients. The following sub-sections estimate the overall as well as the constituent biases in infant and under-five mortality estimates derived using the CEB/CS technique.

3.6.1 Estimating overall bias

The overall bias in estimates derived using the CEB/CS method is determined by comparing the corrected estimates of infant and under-five mortality, denoted by $q^c(1)$ and $q^c(5)$, respectively, obtained after correcting DHS estimates by the Hill and Walker and Hallett *et al* models, with $q^*(1)$ and $q^*(5)$, derived from the standard application of

the CEB/CS at the same points in time. The methods for estimating the various components of the overall bias are described in the following sub-sections.

3.6.1.1 Estimating bias due to the impact of HIV on the mortality of mothers and their children, timing and regression equations

The values of mortality rates obtained from the application of the standard CEB/CS method before conversion to a common index, $q'(x)$, are compared with the $q^c(x)$ values for $x = 1, 2, 3, 5$ and 10 .

In order to enable this comparison, approximations of the values of $q^c(x)$ for ages 2, 3 and 10 have to be made, since the adaptation of the model developed by Hallett *et al* and Hill and Walker only estimates the bias in survey estimates of $q(1)$ and $q(5)$. The hyperbolic model proposed by Blacker and Brass (2005) was fitted to values of $q^c(1)$ and $q^c(5)$ to provide estimates of $q^c(x)$ at other ages. The model is

$$l(x) = (1 + \alpha x)^{-\beta}$$

where x represents the age, and α and β are constants which represent level and pattern of death rates, respectively.

Although, the Blacker and Brass model was designed to capture very early child mortality, the result was extended to allow the estimation of bias in estimates derived from women aged 35-39 at the time of the survey. This extension is not expected to give problems since mortality beyond age 5 is generally low hence it is assumed, without loss of generality, that the bulk of under-10 deaths are reflected in $q(5)$ so the cumulative survival is not affected.

In order to estimate the combined bias due to the impact of HIV on the mortality of the mothers and their children, timing and the regression equation, it is assumed that the bias in the proportion of women dying within a year of the birth of a child is negligible, thus $q^c(1)$ can be approximated by the observed $q(1)$. The size of the bias due to HIV is expected to be small for recent births because women who give rise to these births are likely to be alive to report births and deaths of their children at the time of the survey. The bias is expected to increase as the time between giving births and the survey date increases because of the higher probability of HIV-related maternal deaths.

3.6.1.2 Estimating model life table conversion bias

The CEB/CS technique uses a model life table to convert the estimates of mortality by age group into a common index of mortality over time. The standard application uses a model life table with an age-specific pattern of mortality that is not consistent with the age-specific patterns of mortality found in HIV populations (Ward and Zaba, 2008). The use of such a model life table contributes some error to the estimates derived using the CEB/CS technique. This bias is calculated by comparing the $q(1)$ and $q(5)$ estimates obtained using the Coale-Demeny life tables and the $q(1)$ and $q(5)$ derived using a standard life table allowing for the impact of HIV. In both cases the estimates are derived after correcting the data for bias due to the impact of HIV on the mortality of mothers and their children.

The life table which incorporates the effect of HIV is constructed from a projection of the national populations allowing for the national HIV epidemic. The projection is produced using the Spectrum model and the population is assumed to be closed for migration and applies to the middle of each year. The population is projected in single years and survivorship ratios are estimated for single years of age for each period between consecutive years by dividing population $x+1$ in the year $t+1$ by the population aged x in the year t . This ratio is assumed to be approximately equal to the ratio L_{x+1} / L_x .

The number of births between the middle of year t and the middle of year $t+1$, $B^* = 0.5 \times (B_t + B_{t+1})$. The probability of surviving to $t+1$ of births in the previous year is obtained by dividing the population aged 0 in year $t+1$ by B^* to give an estimate of L_0 / l_0 . A radix of $l_x = 1$ is assumed and the L_x values are obtained by successive substitutions. The l_x values for ages 1-10 years for each year were derived using the model, $l_x = (1 + \alpha x)^{-\beta}$, proposed by Blacker and Brass (2005). The alpha and beta values for the model were chosen by minimising the sum of squared differences between the L_x estimated from the projected population and

$$L_x = \int_x^{x+1} (1 + \alpha \varphi)^{-\beta} d\varphi = \frac{(1 + \alpha(x+1))^{-\beta+1} - (1 + \alpha x)^{-\beta+1}}{\alpha(1 - \beta)}.$$

To compare between the standard life tables, that is the Princeton model life tables and the model life tables incorporating the effect of HIV for Zimbabwe, Kenya, Lesotho, Malawi, Zambia and Namibia, mean absolute errors between the ${}_n q_x$ values

from the observed and the standard life tables were calculated. Time series plots of the mean absolute error were constructed for each country and standard life table used to derive $q(1)$ and $q(5)$ in the CEB/CS method and the model life table with the HIV effect. The model that provides the best fit for the observed data from each country is the one with the smallest mean absolute errors. The following index was used to assess the mean absolute error:

$$\text{Mean absolute error} = \sum \left| \left(1 - \frac{{}_n\hat{q}_x}{{}_nq_x} \right) \right| / n, \text{ where } n \text{ is the number of age groups}$$

in each life table.

The mean absolute errors were calculated for each of the 15-19, 20-24, 25-29, 30-34 and 35-39 year age groups of women. The mean absolute errors were plotted against the time location estimates for the ${}_nq_x$ values for each age group of women from the 20-24 to the 35-39 age groups as shown in Appendix G. The time series plots for the mean absolute errors showed variation between age groups and across countries. The plots for Zimbabwe, Kenya and Zambia showed that the model life table incorporating the effect of HIV fitted the observed data better than the Princeton life tables. However, for Malawi, Namibia and Lesotho, the Princeton life tables had lower mean absolute errors compared to the model life tables with HIV effect.

4 RESULTS

This chapter presents the infant and under-five mortality estimates obtained from applying the direct and indirect methods to data obtained from the 1988, 1994, 1999 and 2005 Demographic and Health Surveys undertaken in Zimbabwe. Section **Error! Reference source not found.** presents the corrected estimates of infant and under-five mortality obtained after adjusting for under-reporting, due to the impact of HIV, using correction factors obtained from the implementation of the Hallett *et al.* and the Hill and Walker procedures. Estimates of under-estimation of rates based on full birth histories calculated from the Hallett *et al* and the Hill and Walker models are also presented. Finally, estimates of various components of bias in estimates derived using the CEB/CS method are calculated. The results for Zimbabwe are presented in full since more information was available to allow verification of results, especially for the Hallett *et al* method. To investigate the behaviour of the bias in other countries with different background mortality rates and at different stages of the HIV epidemic, a summary of the findings from data from the other five African countries is presented in section 4.6.

4.1 Estimation of infant and under-five mortality before correcting for the impact of HIV in Zimbabwe

This section presents estimates of the infant and under-five mortality using data from the 1988, 1994, 1999 and 2005-06 Zimbabwe DHSs. It should be noted that the childhood mortality estimates presented are likely to be downwardly biased due to HIV bias. The bias results from the fact that children born to HIV-positive mothers have an elevated risk of dying but their mothers are less likely to be included in surveys due to illness or death resulting in underreporting of mortality.

4.1.1 Estimation of infant and under-five mortality from complete birth histories

Infant and under-five mortality estimates are calculated using the direct approach using data on the number of children living or deceased provided by women of reproductive age interviewed in the DHSs. Table 4.1 below presents the calculated infant and under-five mortality estimates obtained from the application of the direct estimation methods to the data obtained from the 1988, 1994, 1999 and 2005-06 Zimbabwe DHSs.

The estimates of infant and under-five were calculated for periods 0-4, 5-9, 10-14, 15-19 and 20-24 years before the survey. The reference times are chosen to be the mid-point of each five-year period before the date of the survey which in turn is chosen

as the mid-point of the data collection period of each DHS. The calculated rates are the same as those published in the various DHS reports.

Table 4.1 Direct childhood mortality estimates for Zimbabwe

PERIOD	1988 ZDHS			1994 ZDHS			1999 ZDHS			2005-06 ZDHS		
	TIME	q_0	${}_5q_0$	TIME	q_0	${}_5q_0$	TIME	q_0	${}_5q_0$	TIME	q_0	${}_5q_0$
0-4	1986.4	49.1	70.6	1992.2	52.8	77.1	1997.3	65.0	102.1	2003.3	59.9	82.5
5-9	1981.4	64.4	101.0	1987.2	49.6	74.8	1992.3	53.8	76.9	1998.3	37.1	53.8
10-14	1976.4	54.7	91.8	1982.2	59.6	101.4	1987.3	39.8	59.0	1993.3	40.5	58.2
15-19	1971.4	67.6	113.4	1977.2	63.0	109.9	1982.3	50.8	81.7	1988.3	38.2	56.0
20-24	1966.4	59.5	120.9	1972.2	50.3	94.0	1977.3	57.6	108.6	1983.3	47.7	77.7

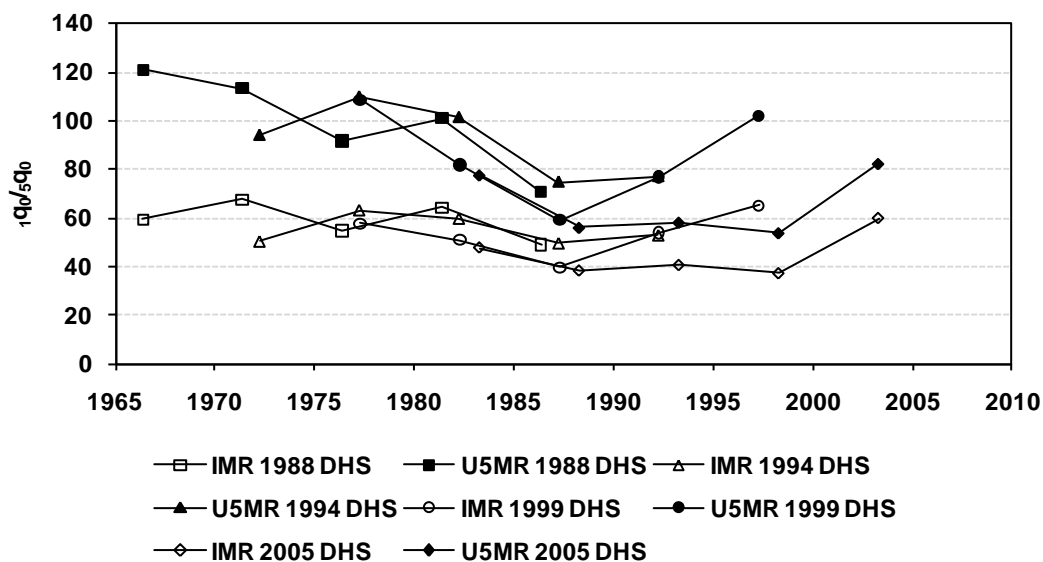
The trends in infant and under-five mortality estimates for the period 1966 to 2003 are shown in Figure 4.1. The general trend in the estimates suggests that mortality declined from the late 1960s until the late 1980s and began to rise around the 1990s. A comparison of the 0-4 year periods for the consecutive DHSs shows that reported infant mortality remained relatively constant at between 49.1 infant deaths per 1000 births in 1986 and 52.8 infant deaths per 1000 births for the 0-4 year period before the 1988 ZDHS. However, there was a marked increase of just over 23% in infant mortality between the 0-4 year periods before the 1988 and 1994 ZDHS, probably because of HIV/AIDS, although the rate decreased from 65 to 60 deaths per 1000 births in the ensuing years up to the 0-4 year period before the 2005-06 ZDHS.

The under-five mortality estimates exhibit a more apparent decline up to the mid-1980s although the same behaviour as that of infant mortality is noted in the trend over time. For the corresponding 0-4 year periods before the four surveys, the first two surveys show an increase of nearly 9.2% to 77.1 deaths per 1000 births between the 1988 and 1994 ZDHSs and an increase of nearly 33% in under-five mortality was observed between the 1994 and 1999 estimates.

A notable feature is the lack of consistency in the mortality estimates for roughly the same reference dates in the 1999 and 2005-06 DHSs. Comparing the infant mortality estimates for the period 0-4 years before the 1999 ZDHS to the infant mortality estimate for the period 5-10 years before the 2005-06 ZDHS gives a difference of around 63%. The most likely explanation is that most women did not survive to report the births and deaths of their children at the time of the 2005-06 DHS as a result of HIV-related deaths or illness. Children born to HIV-positive women have an excess risk of dying so failure to capture their deaths will lead to downwardly biased estimates

of infant and under-five mortality. The same observation is made on the under-five mortality estimates which show a difference of about 33% between the 1997 and 1998 estimates derived from the two DHSs.

Figure 4.1 Infant and under-five mortality from direct estimation; Zimbabwe 1988, 1994, 1999 and 2005 DHSs



4.1.2 Indirect infant and under-five mortality estimates for Zimbabwe

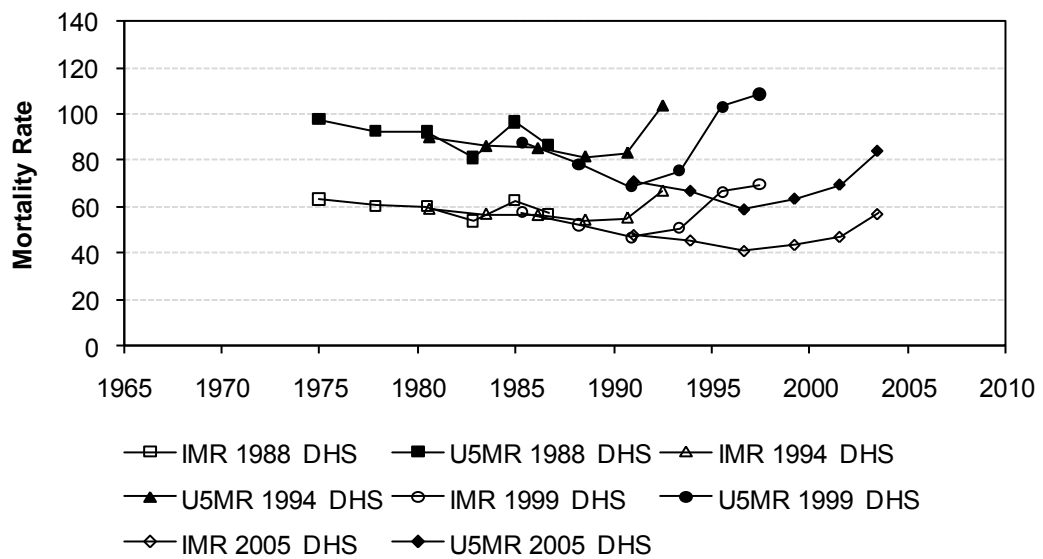
This section presents estimates of infant and under-five mortality derived using the children ever born/children surviving (CEB/CS) method. Partial birth histories data are obtained from the DHSs from the responses provided by women of reproductive ages to questions on the number of children they have ever given birth to and the survival status of those children. The results presented in Table 4.2 exclude information from women aged 15-19 since children of very young mothers are expected to have higher than average mortality thus they are usually ignored in analyses. The estimates shown in the table are for the period 1975 to 2003.

Table 4.2 Indirect childhood mortality estimates for Zimbabwe

1988 DHS			1994 DHS			1999 DHS			2005-06 DHS		
Time	q_0	${}_5q_0$	Time	q_0	${}_5q_0$	Time	q_0	${}_5q_0$	Time	q_0	${}_5q_0$
1986.7	57.0	86.6	1992.5	66.6	103.6	1997.4	69.5	108.6	2003.5	57.0	83.8
1984.9	62.7	96.6	1990.7	55.1	83.2	1995.6	66.4	103.1	2001.5	47.2	69.4
1982.8	53.9	81.2	1988.5	54.2	81.6	1993.3	50.8	75.7	1999.2	43.7	63.4
1980.5	60.3	92.3	1986.1	56.2	85.2	1990.9	46.8	68.8	1996.7	41.1	58.7
1977.9	60.4	92.6	1983.5	56.7	86.1	1988.2	52.3	78.4	1993.9	45.6	67.7
1975.0	63.3	97.7	1980.6	58.9	89.9	1985.3	57.7	87.8	1991.0	48.0	70.8

There has been a general decline in infant mortality over the period considered, as can be seen from Figure 4.2, although the 1994 and 1999 ZDHSs show a slight increase in infant mortality for recent estimates. As with the estimates from complete birth histories, the infant mortality estimates from indirect estimation using the 1999 and 2005-6 DHSs show the same inconsistencies for approximately the same period indicating that there could be problems with at least one of the DHSs.

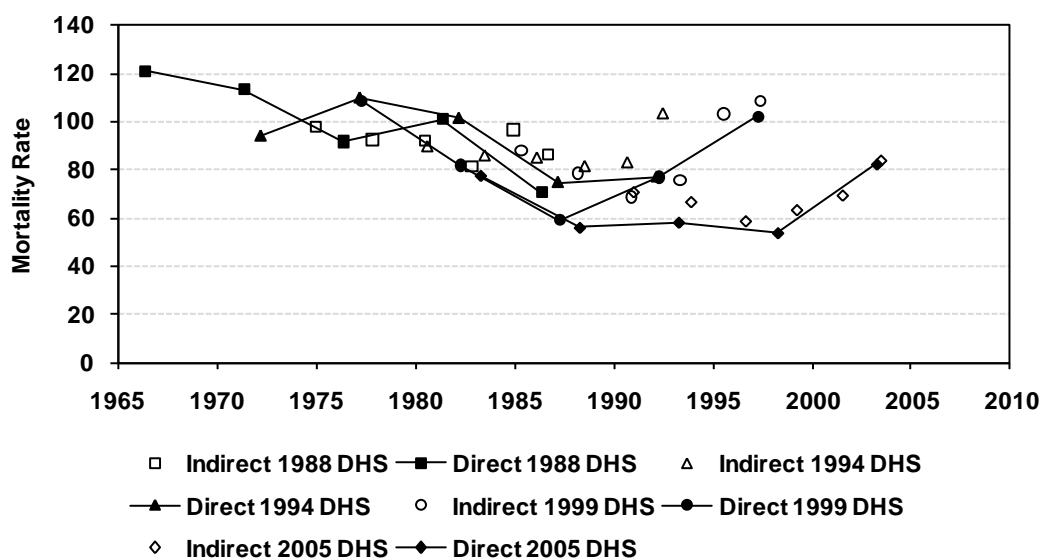
Figure 4.2 Infant mortality and under-five from the CEB/CS method; Zimbabwe 1988, 1994, 1999 and 2005-6 DHSs



4.2 Comparison of trends from direct and indirect estimates

Infant and under-five mortality estimates from direct and indirect estimation methods are presented on the same graph, Figure 4.3, to investigate consistency in trends produced by the two methods. The graph shows that the estimates from indirect methods are higher than the estimates from full birth histories especially for periods closer to survey dates. These discrepancies are, however, more pronounced for under-five estimates from the 1988 and 1994 ZDHSs than in more recent surveys. It should, however, be noted that the estimates from the direct and indirect methods are not directly comparable because of intrinsic differences between the two methods which can lead to faulty conclusions (Nannan *et al.*, 2006; Adetunji, 1996).

Figure 4.3 Direct and indirect estimates of under-five mortality for Zimbabwe



4.3 Estimation of bias in survey estimates of childhood mortality

Estimates of bias as a result of correlation between HIV-related mortality between mothers and their children in survey estimates of infant and under-five mortality are derived from the Hill and Walker workbook and the model developed by Hallett *et al.* The bias estimates derived from the model developed by Hallett *et al.* are used to adjust survey estimates of infant and under-five mortality for the 0-4 year period before each survey while the correction factors from the Hill and Walker model are used to adjust survey estimates for the 0-4, 5-9 and 10-14 year periods before the latest survey. The next sections present the bias estimates from the two methods.

4.3.1 Estimation of HIV bias using the Hallett *et al.* model

The bias estimates shown in Table 4.3 below were produced from the model proposed by Hallett *et al.* The correction factors refer to the period 2.5 years before the 1994, 1999 and 2005-6 ZDHSs. As expected, the percentage of under-reporting for infant mortality is lower than the percentage of under-reporting for under-five mortality. This is because the chances of a woman having died due to HIV/AIDS by the time of the survey increase with time before the survey. Women who are supposed to report under-five mortality are more likely to have died by the time of the survey compared to women who report infant mortality.

The bias in child mortality estimates is expected to increase with time as the epidemic evolves and more women become infected although this will also depend on other factors such as intervention and treatment. Table 4.3 shows that the estimates of bias in under-five mortality increased by close to 10 percentage points from 0.9% in

1992 to 10.7% in 2003. The model results also show that bias in infant mortality increased from 0.7% in 1992 to 8.5% in 2003.

Table 4.3 Estimates of under-reporting (%) from the Hallett *et al.* and Hill and Walker models for five years preceding the survey

Year	Infant Mortality		Under-five Mortality	
	Hallett <i>et al.</i>	Hill and Walker	Hallett <i>et al.</i>	Hill and Walker
2003	8.5	6.8	10.7	11.5
1997	3.6	6.6	4.5	10.1
1992	0.7	3.9	0.9	5.0

4.3.2 Estimation of HIV bias using the Hill and Walker model

The bias estimates for the 0-4, 5-9 and 10-14 year periods before the 2005-6 ZDHS are shown in Table 4.4 below. The estimates of bias for infant and under-five mortality are based on pre-AIDS mortality estimates of 50 and 75 deaths per 1000 births for infant and under-five mortality respectively, which were obtained from the 1988 ZDHS. The percentage of under-reporting increase going back in time with estimated under-reporting for infant mortality of 7% in 2003 and 28% in 1993. Under-five decreased by 16 percentage points from 28% in 1993 to 12% in 2003 although an increase of 1 percentage point was observed for the period from 1998 to 1993.

Table 4.4 Correction factors from the Hill and Walker model

Estimates for single survey				Estimates for multiple surveys			
Period	Reference Time	Infant Mortality	Under-five Mortality	Period	Reference Time	Infant Mortality	Under-five Mortality
0-4	2003	6.8	11.5	0-4	2003	6.8	11.5
5-9	1997	24.3	29.1	0-4	1997	6.6	10.1
10-14	1993	28.2	28.0	0-4	1992	3.9	5.0

Estimates of bias from the Hill and Walker method were calculated for each of the 0-4 year periods before the 1994, 1999 and 2005-06 ZDHSs to enable the comparison with the bias estimates from the Hallett *et al.* method and the results are presented in Table 4.3. The estimates for the 0-4 year period before the 2005-06 DHS from the two methods are reasonable close but the Hill and Walker estimates are consistently higher than those from the Hallett *et al.* method for the earlier periods for both infant and under-five mortality. The bias estimates for the 0-4 year period before the 1999 DHS from Hill and Walker exceed those from the Hallett *et al.* method by 45 and 55 per cent, respectively for both infant and under-five mortality while the difference is much higher at 82 per cent for both infant and under-five mortality for the

0-4 year period before the 1994 DHS. The causes of the differences in estimates of bias from the two methods were not found despite some investigations.

4.4 Comparison between estimates of infant and under-five mortality corrected using the Hallett *et al* and Hill and Walker methods

below shows estimates of infant and under-five mortality corrected for underestimation using correction factors derived from both the Hill and Walker and the Hallett *et al.* methods. The estimates corrected using the Hill and Walker method apply to the middle of the 0-4, 5-9 and 10-14 year periods before the 2005-06 DHS. The estimates from the 1994 and 1999 and 2005-06 DHSs used in the Hallett *et al.* method apply to the middle of the 0-4 year period before each survey. The corrected estimates from both estimates are close at the start and end of the period under consideration but show a marked difference at the middle dates around 1997 and 1998.

The contrasting behaviour of the trend in estimates over time from the two methods is of concern since it shows a marked difference of about 19% in infant mortality estimates and around 29% in under-five mortality estimates within a year, from around 1997 and around 1998. This is probably a result of the difference in data quality between the 1999 and the 2005-06 ZDHSs. The Hill and Walker correction factors apply to a single DHS while the Hallett *et al.* method needs three DHSs to produce comparable estimates. This means that consistency between corrected estimates produced by the two methods can only be achieved if successive DHS are consistent.

The corrected estimates were also calculated for the other countries in the study and the results are presented in Appendix A. The corrected estimates produced from the Hill and Walker methods were first adjusted for the problem of birth transference. The results show that for countries whose successive DHSs show plausible consistency in estimates of infant and under-five mortality, the two methods produce reasonably close estimates. The corrected estimates from the Hallett *et al.* method, however, tend to be greater than those from the Hill and Walker method for periods further away from the survey date.

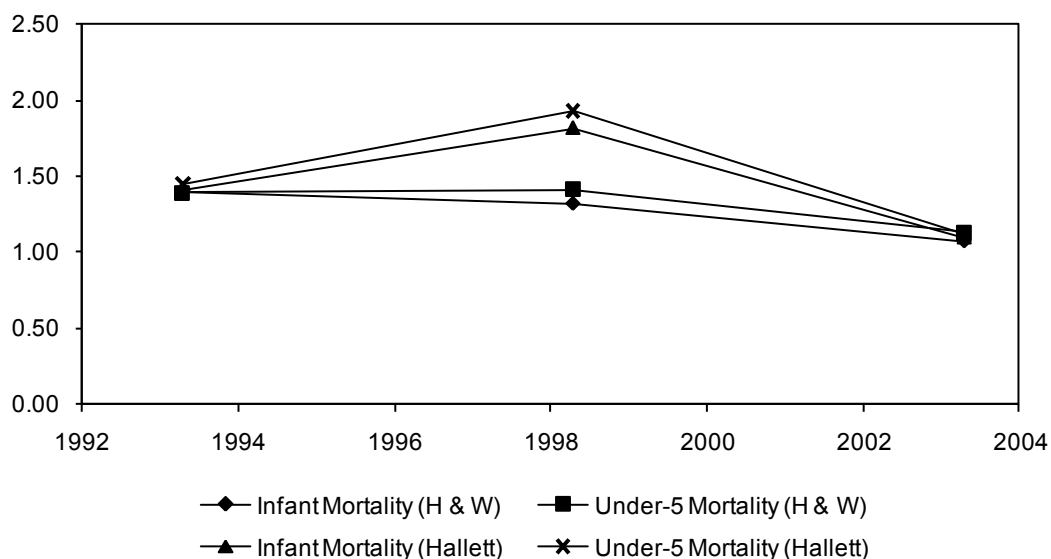
Table 4.5 Comparison of estimates of infant and under-five mortality corrected using correction factors from Hallett *et al.* and Hill and Walker methods

Infant Mortality				Under-five Mortality			
TIME	Hill & Walker	TIME	Hallett <i>et al</i>	TIME	Hill & Walker	TIME	Hallett <i>et al</i>
2003	64.4	2003	65.6	2003	92.7	2003	91.8
1998	48.9	1997	67.4	1998	76.2	1997	106.9
1993	55.7	1992	53.6	1993	80.5	1992	77.8

The corrected estimates from the two methods were also compared to the infant and under-five mortality estimates derived directly from full birth history data from the 2005-06 ZDHS. Figure 4.4 shows the ratios of the estimates of infant and under-five mortality derived from the Hill and Walker method and the Hallett *et al* to the estimates of infant and under-five mortality derived from the 2005-06 ZDHS. The corrected estimates from the Hallett *et al* method were used to produce linear interpolated estimates to match the DHS dates to enable comparison with the reported 2005-06 DHS estimates.

The ratios of corrected estimates of infant and under-five mortality from the two methods to the DHS estimates are reasonably close in cases where successive DHSs produce matching estimates of infant and under-five mortality for the same reference times. In this case, the ratios are almost the same in 1993 and 2003 but very different for the period around 1998. The marked difference between the ratios for both infant and under-five mortality around 1998 is probably the result of inconsistency between the 1999 and 2005-06 ZDHSs. This can be shown by considering the infant and under-five mortality estimates that apply at almost the same reference date obtained from the two surveys. The corrected estimates from the Hallett *et al.* method are sensitive to such inconsistencies since they are derived from estimates applying to the 0-4 year period before each of the three successive surveys.

Figure 4.4 Ratios of estimates of infant and under-five mortality corrected using correction factors from Hallett *et al.*, Hill and Walker methods to the DHS estimates



4.5 Estimation of overall bias in estimates derived using the CEB/CS method

Since the corrected estimates of infant and under-five mortality from Hallett *et al* and Hill and Walker methods are similar for countries whose successive DHS surveys give matching estimates of infant and under-five mortality for similar reference times such as Kenya, Lesotho, Malawi and Zambia, it was decided to calculate the biases using the Hill and Walker method alone for all countries. In addition, since the CEB/CS estimates of infant and under-five mortality used in estimating bias come from a single DHS, using the Hill and Walker method will help to minimise errors from other sources such as sampling errors associated with using several DHSs. The results for Zimbabwe will be presented in this section and some of the results for the other countries will be presented in section 4.6.

Table 4.6 shows the overall bias in infant and under-five mortality estimates derived using the CEB/CS method. Since it might be expected that the overall relative bias would be the same for infant and under-five mortality estimates because the estimates are derived from reports of the same age groups of women, the overall bias by the age group of the mother is estimated as an average of the two sets of estimates. The estimates of infant and under-five mortality obtained after adjusting DHS estimates with the correction factors from the Hill and Walker method apply to the date at the middle

of each of the 0-4, 5-9 and 10-14 year period before the survey². The $q^c(1)$ and $q^c(5)$ values were obtained by interpolating between corrected rates for the 0-4, 5-9 and 10-14 years period before the survey. The negative sign on the bias estimate means that the bias leads to an underestimation of mortality.

The overall biases in the estimates derived using the CEB/CS method are obtained by comparing the corrected estimates of infant and under-five mortality, $q^c(1)$ and $q^c(5)$, obtained after correcting the 2005-06 ZDHS estimates of bias due to HIV with estimates of infant and under-five mortality, $q^*(1)$ and $q^*(5)$, obtained from the usual application of the CEB/CS estimates at the same points in time.

The results show, ignoring the bias in estimates derived from women aged 15-19 because of a lack of representativeness and high proportion of first births, which lead to an elevated risk of mortality in their children, that the CEB/CS method underestimates the mortality rates by between 11% and 22%. The bias in estimates drawn from women in the 20-24 age group is noticeably lower than that of older age groups of women, which is reasonably level, allowing for random fluctuation. An explanation for this is that prevalence would not be at peak and most women in this age group would still be alive to report on the mortality of their children at the time of the survey even if they were infected by the time of giving birth.

Table 4.6 Overall bias in infant and under-five mortality estimates derived from the CEB/CS method for the 2005-06 ZDHS

Age Group	Time	$q^*(1)$	$q^c(1)$	Bias in $q(1)$	$q^*(5)$	$q^c(5)$	Bias in $q(5)$	Overall bias
20-24	2003.5	57.0	65.0	-12%	83.8	93.3	-10%	-11%
25-29	2001.5	47.2	59.0	-20%	69.4	86.9	-20%	-20%
30-34	1999.2	43.7	51.8	-16%	63.4	79.3	-20%	-18%
35-39	1996.7	41.1	51.1	-20%	58.7	77.6	-24%	-22%

4.5.1 Estimation of aggregate bias due to the impact of HIV on mortality of mothers and their children, regression coefficients and timing

Table 4.7 presents the combined bias due to the impact of HIV on the mortality of mothers and their children, the use of the regression coefficients and timing. The bias was calculated by comparing the values of mortality rates obtained from the application

² Infant and under-five mortality estimates for Kenya, Lesotho, Namibia, Malawi and Zambia were calculated for the 1-5, 6-10 and 11-15 year periods before the survey to adjust for the effect of birth transference.

of the standard CEB/CS method before conversion to a common index, $q'(x)$, with the $q^c(x)$ values for $x = 1, 2, 3, 5$ and 10 . The $q^c(x)$ values were obtained after correcting for HIV-related bias in infant and under-five mortality estimates derived from the 2005-06 ZDHS using the correction factors obtained from the Hill and Walker model.

The results show, ignoring the first age group, that these biases combined lead to an underestimate of $q(2)$ of 15% and underestimates of 20-25% of $q(x)$ at the older ages clearly contributing significantly to the overall bias.

Table 4.7 HIV non-survivor and regression equation bias, ZDHS 2005-06

Age Group	Time	x	$q'(x)$	$q^c(x)$	Bias
20-24	2003.5	2	66.0	77.2	-15%
25-29	2001.5	3	61.0	78.1	-22%
30-34	1999.2	5	63.4	79.3	-20%
35-39	1996.7	10	67.0	89.0	-25%

4.5.2 Model life table conversion bias

Table 4.8 shows the conversion bias due to the choice of model life table after applying life tables with and without the impact of HIV to convert the $q'(x)$ s to common indices of infant and under-five mortality. The Coale-Demeny North model life table level 19 for combined sexes with an assumed sex ratio at birth of 1.02 was chosen to represent the age pattern of mortality for Zimbabwe. Level 19 was chosen on the basis that it gave a $q(5)$ value which is close to the observed $q(5)$. The standard life table with impact of HIV for the year 2003 was chosen on the assumption that it represented the mortality of children under the age of five years.

The $q'(x)$ values are the $q(x)$ values derived from the CEB/CS method corrected for bias due to the correlation of mortality between HIV-positive mothers and their children, the regression coefficients used and the estimate of timing. The model life table conversion bias is then calculated by comparing the $q(1)$ and $q(5)$ estimates obtained using the Coale-Demeny life tables and the $q(1)$ and $q(5)$ derived using a standard life table allowing for the impact of HIV.

This bias presented in the last two columns of Table 4.8 shows that the choice of model life table affects estimates of infant and under-five mortality only marginally.

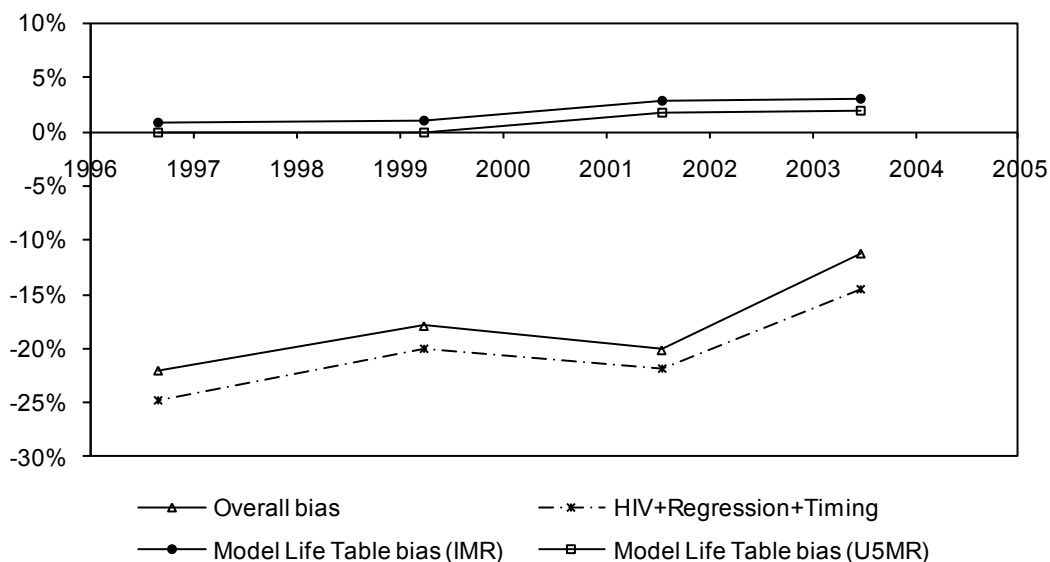
Table 4.8 Estimates of model life table conversion bias for Zimbabwe

Age Group	Date	Coale-Demeny North		Life Table with HIV impact		Bias	
		Infant Mortality	Under-five Mortality	Infant Mortality	Under-five Mortality	Infant Mortality	Under-five Mortality
20-24	2003	65.3	97.5	63.3	95.6	3%	2%
25-29	2002	59.5	89.1	57.8	87.5	3%	2%
30-34	1999	52.8	79.3	52.2	79.3	1%	0%
35-39	1997	51.5	77.4	51.0	77.5	1%	0%

4.5.3 Comparison between overall bias and sum of constituent biases

Figure 4.5 below shows the comparison of the overall bias and the sum of individual biases due to the impact of HIV on the correlation of mortality of mothers and their children, use of regression coefficients, timing and model life table conversion. The influence of the combined bias of the impact of HIV on mortality of mothers and their children, timing and the regression coefficients is partly counteracted by the model life table conversion bias, which inflates the mortality estimates by about 3%. The individual biases seem to be explaining the overall bias quite well allowing for a level of random fluctuation.

Figure 4.5 Comparison of biases in estimates of infant and under-five mortality derived using the CEB/CS method



4.6 Comparison of bias estimates among countries with generalised HIV epidemics

The same procedure was used to estimate bias in infant and under-five mortality in five other African countries with generalised HIV epidemics, which are Kenya, Namibia, Lesotho, Malawi and Zambia.

Table 4.9 shows the estimates of overall bias in estimates of infant and under-five mortality derived from the CEB/CS method. The estimates are presented for the 20-24, 25-29, 30-34 and 35-39 age groups of women and for each of the countries under study. As with the estimates from Zimbabwe, these results also show substantial bias in indirect estimates of infant and under-five mortality. The results also show that, in most cases, larger biases are observed in estimates derived from older women. This is to be expected since the reporting bias is higher in older women because of greater cumulative HIV infection and HIV/AIDS deaths. The overall bias in infant and under-five mortality estimates derived from the CEB/CS method for the countries shown ranges from 3% to 34% for age groups from 25-29 to 35-39 years.

Most countries, with the exception of Namibia and a few fluctuations, show the same trend of increase with age in overall bias estimates from the 25-29 years age group onwards, the 20-24 years age group can be ignored because of high proportion of first births, which may distort the indirect estimates. One striking observation is the large biases observed in Kenya despite the fact that it has relatively lower HIV prevalence rates compared to the other countries.

Countries with higher background infant and under-five mortality have relatively lower biases, as is the case for Malawi and Zambia, compared to the other countries. The explanation for smaller biases in countries with higher background mortality is that the proportion of HIV-related deaths among children will be lower.

Table 4.9 Overall bias (percent) by age group of women and country

Age Group	Country				
	Kenya	Lesotho	Malawi	Namibia	Zambia
20-24	42	-27	11	-1	25
25-29	-4	-7	-4	-21	-3
30-34	-29	-30	-10	-19	-16
35-39	-32	-16	-16	-34	-20

Table 4.10 shows the combined bias due to the impact of HIV on the mortality of mothers and their children, the use of the regression coefficients and timing. This bias contributes significantly to the overall bias hence the pattern of this bias is almost the same as that of the overall bias.

The general pattern observed, with the exception of fluctuations in Namibia, is that the magnitude of the bias due to the impact of HIV on the mortality of mothers and their children, the use of the regression coefficients and timing increases with age as

observed in the estimates for women aged 25-39 years. This is expected since the number of unreported deaths of children born to infected women increases over time since mortality of infected mothers increase with duration since infection. Four of the five countries show a negative bias of between 5% and 6% in the 25-29 year age group indicating that the risk of dying within few years of giving birth is low even if the mother is infected. This is supported by the estimates of survival since infection which are approximated to be around ten years (Artzrouni and Zaba, 2003).

Again the estimates for women in the age group 20-24 show different behaviours in the different countries, with 4 out of 5 countries showing positive biases in this age category. This could be due to the effect of the methodological differences between the CEB/CS method and the direct method for deriving child mortality estimates using full birth history data. The CEB/CS method produces child mortality estimates that are higher than those obtained directly from full birth history data especially for women in younger age groups as observed in others studies (Adetunji, 1996).

Table 4.10 HIV non-survivor and regression equation bias (percent) by age group and country

Age Group	Country				
	Kenya	Lesotho	Malawi	Namibia	Zambia
20-24	39	-26	10	1	23
25-29	-5	-6	-6	-22	-6
30-34	-29	-28	-13	-19	-19
35-39	-31	-14	-20	-34	-26

Table 4.11 shows the bias which is due to the choice of model life table calculated for the countries. The results confirm the observation made in the Zimbabwean case that a small bias is introduced through the use of a model life tables which do not have the impact of HIV. With the exception of Malawi, the observed conversion bias ranges between 0 and +/-3%. This bias is in the margin of error for the application of the CEB/CS method in the first place.

However, the bias in infant mortality (shown in Appendix C) is a bit higher, partly because of the use of estimates of women age 15-19 years in calculating the implied infant mortality estimates for each age group of women. The other explanation is that the indirect estimates of infant mortality were derived from the estimates of under-five mortality, which means that most of the model life tables could not define

the estimates of infant mortality very well since they were chosen based on the under-five mortality estimates.

Table 4.11 Conversion bias (percent) in under-five mortality by age group of women and country

Age Group	Country				
	Kenya	Lesotho	Malawi	Namibia	Zambia
20-24	-2	-3	8	1	-1
25-29	-1	-2	4	1	0
30-34	0	0	0	0	0
35-39	2	3	3	1	1

Although the comparison gives useful validation of results to enable generalisations, it should be noted that the discrepancies brought about by different times in which DHS surveys were held might, to some extent, lead to a distorted assessment. The bias behaves differently at different stages of the epidemic thus comparing countries at different stages of the HIV epidemic can distort the assessment, Hallett *et al* (2010) showed that the bias increases in increasing or stable epidemics.

For example, the latest DHS used for Malawi was conducted in 2004 whereas Lesotho has a DHS survey undertaken in the period 2009-2010. This discrepancy will affect comparisons between these two countries since in 2004 the epidemic in Malawi was likely to be in its early stages compared to the stage of the epidemic in Lesotho in 2009-10. A greater proportion of women in Lesotho are likely to have died due to HIV/AIDS before the 2009-10 DHS compared to the proportion who died of HIV/AIDS in Malawi before the 2004 DHS as a direct consequence of the natural progression of the disease in humans.

Comparisons across countries can also be affected by the differences in the level of background mortality. Countries, such as Lesotho and Zimbabwe, with relatively low background infant and under-five mortality and high HIV prevalence levels have higher biases. Conversely, countries with lower HIV prevalence levels and high background mortality have lower biases because very few women will miss the survey due to HIV-related deaths. The scale of intervention and treatment will also hinder comparisons of bias estimates across countries.

5 DISCUSSION AND CONCLUSION

5.1 Introduction

The overall aim of this research was to estimate the extent of bias in infant and child mortality estimates derived from the Brass CEB/CS due to HIV. This was done by first calculating the best estimates of infant and child mortality derived from data on full birth history adjusting for the impact of HIV using a method developed by Hill and Walker, and then comparing these rates with estimates of infant and child mortality derived using the CEB/CS method. This chapter seeks to explain how the results support the original aims and how the results compare with findings of other related studies. The limitations, scope for further research and conclusion of the research are discussed in later sections.

Due to the lack of reliable vital registration systems in most sub-Saharan African countries, infant and child mortality estimates are often calculated from cross-sectional surveys and national census data. The CEB/CS method is an attractive method for childhood mortality estimation in countries with deficient vital registration systems, due to the short and simple nature of the questions. The succinct nature of the questions helps to minimise errors associated with mortality data and allows the questions to be incorporated in censuses and other large surveys for the collection of childhood mortality measures from population subgroups (Hill, 1991). This study was motivated by the need to adapt this useful method which has been found wanting in the wake of HIV/AIDS, which has led to the suspected violation of most fundamental assumptions that underlie the use of the method leading mostly to downwardly biased estimates (Ward and Zaba, 2008; Mahy, 2003).

To estimate the bias, there was the need initially to find the most accurate estimates of infant and under-five mortality in the six countries involved in this study. These estimates of mortality were obtained by adjusting the infant and under-five mortality estimates from the latest DHS surveys in each country for underestimation due to HIV selection bias. The method proposed by Hill and Walker was used to derive the correction factors for adjusting the DHS estimates because the method involves the use of child mortality estimates from a single DHS, thus obviating the need for several equally successful past surveys. The other reason is that this method is also being used internationally, by the IGME, to correct for HIV bias in child mortality estimates derived from full birth history. Another method proposed by Hallett *et al.* was also

considered and deemed to produce similar results as the Hill and Walker method except in two of the countries, Zimbabwe and Namibia. The confirmation, as shown in most countries, by the two methods gave some assurance with regards to the reliability of the corrected direct estimates of infant and under-five mortality. Although the method proposed by Hallett *et al* applies for the infant and under-five mortality estimates for the period 0-4 years before the survey, estimates for the period 1-5 years before the survey were also calculated and are not significantly different from the original estimates although they were found to be closer to the Hill and Walker estimates.

The results show that HIV causes substantial bias in the estimates of infant and child mortality derived using the CEB/CS method. The bias is more pronounced in estimates derived from the reports of women aged 25 years and older, which is to be expected in a mature HIV epidemic since these are the ages at which HIV mortality is highest. Thus, depending on the progression of the virus and treatment, effects will start showing around five years after infection when the women begin to succumb to HIV-related diseases or even deaths leading to exclusions from surveys. The bias observed in the women aged between 30 and 39 years in the countries considered was shown to be between 10% and 34%.

Mainly, the reporting bias is caused by non-survival of mothers due to HIV related deaths, which in turn depend largely on the HIV prevalence levels, stage of the epidemic, background mortality and extent of intervention efforts in a country. Although the research was not able to quantify all the biases that result from the use of the CEB/CS method, it managed to show that there is not much bias resulting from the use of the model life tables in estimating under-five mortality. Five of the six countries showed bias due to choice of model life table in the range -3% to +3% in under-five mortality estimates. The other country, Malawi, had much higher bias estimate (8%) in the 20-24 age group possibly because the North family of the Princeton model life tables was a poor fit to the non-AIDS mortality in Malawi. This shows that not much bias is introduced by the use of model life tables which don't allow for the impact of HIV in estimating under-five mortality. However, since the life tables allowing for the impact of HIV used in this assessment pertain to the period in the early 2000s, characterised by relatively lower treatment levels, the bias may behave differently when the epidemics evolve and treatment levels change. The changes in these factors will pose additional problems, since these factors may lead to changes in age patterns of mortality resulting in a different behaviour of the model life table bias.

The estimates of overall bias derived from reports by women in the 20-24 year age group are positive in some of the countries. This indicates that the mortality estimates from the CEB/CS method for children born to these women are higher than the corrected estimates of mortality corresponding to this age group of women. The high indirect estimates could be a result of a high proportion of children ever born to these women being born when they were young (under age 20); this would also explain the discrepancy between the indirect estimates and the direct estimates used in the derivation of corrected child mortality rates. Other studies have also shown that indirect estimates can be higher than direct estimates, especially for estimates close to the survey dates (Hill, 1991; Adetunji, 1996). Hill (1991) indicated that the differences in direct and indirect estimates of child mortality from the same data source may be a result of the manner in which the two methods allocate deaths and exposure to risk for specific time periods.

5.2 Limitations of the study

One of the most important limitations of this study is that the estimates of infant and under-five mortality derived from the direct and indirect methods are often not the same, irrespective of the impact of HIV/AIDS. Attempts to establish the bias in estimates in the manner used in the study will be confounded by these inherent differences between the methods. This is especially so for estimates derived from women in younger age groups. As postulated by Hill (1991), the indirect estimates from younger age groups of women tend to exceed the estimates from direct estimates for the recent period especially if changes in mortality have not been linear.

Although the direct estimates from full birth history data are thought to be more reliable than the indirect estimates, there is not enough evidence to validate this claim since the true level of infant and under-five mortality is not known (Adetunji, 1996). Thus, the validity of the results of this study depends on the accuracy of the direct estimates calculated from the DHS data used in estimating bias.

The accuracy of estimates of bias depends a lot on the reliability of the assumptions used in the process of deriving correction factors that are used to adjust the DHS estimates for underestimation due to HIV. One such assumption is that the HIV prevalence level in pregnant women is representative of the HIV prevalence in the entire population. Since the correction factors make use of HIV prevalence levels in pregnant women, they will only remain reliable if there is no selection bias in women who become pregnant. Other studies have indicated that HIV-positive women who

have been infected for some time have a lower risk of falling pregnant meaning that there is also a bias in the HIV prevalence estimates since a proportion of these usually older HIV-positive women do not visit the ante-natal clinics where the HIV data are captured (Hallett *et al.*, 2010; Artzrouni and Zaba, 2003). The reliability of the HIV prevalence and, hence, the correction factors therefore depends on the extent of this bias.

The Hill and Walker method assumes that the AIDS-free mortality is constant. If background mortality is decreasing, then bias due to HIV will be underestimated since the actual fraction of HIV-related deaths among children would be higher than the calculated value. However, the effect is likely to be the small.

The Hill and Walker method also assumes that children who become infected at birth have a 62.5% chance of dying by age 5 and that ARV treatment has no effect on the mortality risk for the period. The mortality schedule used to estimate this probability was derived from cohort studies. This means that the accuracy of bias estimates will depend on the representativeness of the mortality schedule to other populations. Improvements in coverage of PMTCT would result in a lower risk of dying than the assumed probability leading to reduction in estimates of bias. However, violation of this assumption is likely to have minimal impact because of the relatively low fertility among HIV-positive women.

5.3 Scope for future research

It is important to use data from other countries to assess the behaviour of the bias due to HIV in child mortality estimates derived using the CEB/CS method. In light of biases resulting from the methodological differences in the two methods of estimating infant and under-five mortality, it is critical to consider a 'true' level of infant and under-five mortality estimated from other methods so that the true bias due to HIV in estimates derived using the CEB/CS method can be ascertained. It is also imperative to find ways of isolating the effect of background mortality, as different levels of background mortality affect comparisons of the effect of HIV across countries and within the same country at different times. There is also the need to assess the impact of fertility on the behaviour of the bias since it is likely that women will be infected at different ages when they are still potentially capable of bearing children.

The study was not able to find correction factors for adjusting estimates of infant and under-five mortality according to prevalence levels of HIV. There is a need for

further study to establish this link which can help in the correction of survey estimates of mortality.

5.4 Conclusion

This research set out to find the extent of bias in estimates of infant and under-five mortality. The Hallett *et al* and Hill and Walker models were chosen to provide correction factors for adjusting infant and under-five mortality estimates from the latest DHSs in the selected countries for under-reporting due to the effect of HIV. Although the two methods provide relatively similar estimates of under-reporting for the 0-4 year period before the latest DHSs, they do not agree for the other estimates produced for earlier periods. Mortality estimates adjusted for HIV-bias using correction factors from the Hill and Walker method were eventually chosen for the estimation of bias in estimates derived from the CEB/CS method.

The findings show that most of the bias in infant and under-five mortality is due to HIV non-survival, which also affects direct estimates. It was also shown that there was insignificant bias resulting from the use of the Coale-Demeny model life tables used in the Brass CEB/CS method. This means that with improvement in intervention and treatment in particular the situation will be more or less the same as there was no HIV implying that the CEB/CS method can be used, without adjustment, without much loss of accuracy.

The other observation is that the percentage bias in estimates derived from the direct estimates is almost similar to, if not larger than, that in indirect estimates. For example, in Zimbabwe the direct estimates of U5MR for the periods 5-9 and 10-14 years before the survey are underestimated by around 29% and 28% respectively. Comparing this to the overall bias estimates in age groups between 25 and 39 years shows that, in most cases, indirect estimates have lower bias. The explanation for this may be that a proportion of the HIV bias is counteracted by the inherent bias in the indirect method since, even in the absence of HIV, it appears to produce mortality estimates that are higher than the mortality estimates that are derived directly from full birth history data.

REFERENCES

- Adetunji, J. 2000. "Trends in under-5 mortality rates and the HIV/AIDS epidemic", *Bulletin of the World Health Organisation*, **78** (10): 1200-1206.
- Adetunji, J. and Bos, E. R. 2006. "Levels and trends in mortality in sub-Saharan Africa: an Overview" in Jamison, D. T., Feachem, R. D., Makgoba, M. W., Bos, E. R., Baingana, F. K., Hofman, K. J. & Rogo, K. O. (eds). *Disease and Mortality in sub-Saharan Africa*. Second Edition ed. Washington, D. C., The World Bank, 11-14.
- Adetunji, J. A. 1996. "Infant mortality levels in Africa: Does method of estimation matter?" *GENUS*, **LII** (3-4): 89-106.
- Ahmad, O. B., Lopez, A. D. and Inoue, M. 2000. "The decline in child mortality: a reappraisal", *Bulletin of the World Health Organisation*, **78** (10): 1175-1191.
- Artzrouni, M. and Zaba, B. 2003. "HIV-induced bias in the estimation of child mortality using birth history reports," Presented at the technical meeting on HIV and Child mortality. London, March 2003.
- Bicego, G. and Ahmad, O. B. 1996. *Infant and child mortality*. DHS Comparative Studies No. 20. Calverton, Maryland: Macro International Inc.
- Blacker, J. and Brass, W. 2005. "The estimation of infant mortality from proportions dying among births in the last 24 months", *Southern Africa Journal of Demography*, **10** (1&2): 25-42.
- Brass, W., Coale, A., Demeny, P., *et al.* 1968. *The Demography of Tropical Africa*. Princeton, New Jersey: Princeton University Press.
- Central Statistical Office [Zimbabwe] and Macro International Inc. 2007. *Zimbabwe Demographic and Health Survey 2005-06*. Calverton, Maryland: CSO and Macro International Inc.
- Coale, A. and Demeny, P. 1983. *Regional model life tables and stable populations*. New York: Academic Press.
- Coale, A. J. and Trussell, J. 1974. "Model fertility schedules: variations in the age structure of childbearing in human populations", *Population Index*, **40** (2): 185-258.
- Curtis, S. L. 1995. *Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys*. Occasional Papers No. 3. Calverton, Maryland: Macro International Inc.
- Feeney, G. 1980. "Estimating infant mortality trends from child survivorship data ", *Population Studies*, **34** (1): 109-128.

- Garenne, M. and Gakusi, E. 2006. "Health transitions in sub-Saharan africa: overview of mortality trends in children under 5 years old (1950-2000) ", *Bulletin of the World Health Organisation* **84** (6): 470-478
- Hallett, T. B., Gregson, S., Kurwa, F., *et al.* 2010. "Measuring and correcting biased child mortality statistics in countries with generalized epidemics of HIV infection", *Bulletin of the World Health Organisation*, **88** (10): 761-768.
- Hill, K. 1991. "Approaches to the measurement of childhood mortality: A Comparative Review", *Population Index* **57** (3): 368-382.
- Inter-agency Group for Child Mortality Estimation. 2010a. *Child mortality estimates*. <http://www.childmortality.org/cmeMain.html>. Accessed: 13 March 2011.
- Inter-agency Group for Child Mortality Estimation. 2010b. *Estimation Methods Used by the UN Inter-agency Group for Child Mortality Estimation*. http://www.childinfo.org/stock/documents/methods_for_Estimating_Child_Mortality_2010.pdf. Accessed:17 January 2011.
- Knodel, J. and Chamrathirong, A. 1978. "Infant and child mortality in Thailand: levels, trends and differentials as derived through indirect estimation techniques." *Papers of the East-West Population Institute*, **57**
- Korenromp, E. L., Arnold, F., Williams, B. G., *et al.* 2004. "Monitoring trends in under-5 mortality rates through national birth history surveys", *International Journal of Epidemiology*, **33** (6): 1293-1301.
- Lallement, C., Halembokaka, G., Baty, G., *et al.* 2010. "Impact of HIV/AIDS on child mortality before highly active antiretroviral therapy era: A study in Pointe-Noire, Republic of Congo", *Journal of Tropical Medicine*, **10.1155** 6.
- Mahy, M. 2003. "Measuring child mortality in AIDS-affected countries," Workshop on HIV/AIDS and adult mortality in developing countries. New York, USA 8-13 September 2003.
- Marston, M., Zaba, B., Salomon, J. A., *et al.* 2005. "Estimating the net effect of HIV on child mortality in African populations affected by generalised HIV epidemics", *Journal of Acquired Immune Deficiency Syndromes*, **38** (2): 219-227.
- Nannan, N., Timaeus, I. M., Laubscher, R. and Bradshaw, D. 2006. "Levels and differentials in childhood mortality in South Africa", *J.Biosoc.Sci*, **2007** (39): 613-632.
- Ng'weshemi, J. Z. L., Urassa, M., Isingo, R., *et al.* 2003. "HIV impact on mother and child mortality in rural Tanzania", *Journal of Acquired Immune Deficiency Syndromes*, **33** (3): 393-404.
- Preston, S. H., Heuveline, P. and Guillot, M. 2001. *Demography: Measuring and modelling population processes*. Blackwell Publishing.
- Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., *et al.* 2010a. "Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic

- analysis of progress towards Millenium Development Goal 4", *Lancet*, **375** (9730): 1988-2008.
- Rajaratnam, J. K., Tran, L. N., Lopez, A. D. and Murray, C. J. L. 2010b. "Measuring under-five mortality: Validation of new low-cost methods", *PLoS Med*, **7** (4):
- Rutstein, S. O. and Rojas, G. 2003. *Guide to DHS Statistics*. Demographic and Health surveys. Calverton, Maryland: ORC Macro.
- Sullivan, J. M. 1972. "Models for the estimation of the probability of dying between birth and exact ages of early childhood", *Population Studies*, **26** (1): 79-98.
- Timaeus, I. M. 1998. "Impact of the HIV epidemic on mortality in sub-Saharan Africa: evidence from national surveys and censuses", *Journal of Acquired Immune Deficiency Syndromes*, **12** (Suppl 1): S15-S27.
- Trussell, J. 1975. "A Re-estimation of the multiplying factors for the Brass technique for determining childhood survivorship rates", *Population Studies*, **29** (1): 97-108.
- UNICEF. 2010. *Estimation Methods Used by the UN Inter-agency Group for Child Mortality Estimation*.
http://www.childinfo.org/stock/documents/methods_for_Estimating_Child_Mortality_2010.pdf. Accessed:17 January 2011.
- United Nations. 1983. *Indirect techniques for demographic estimation: Manual X*. New York: United Nations.
- United Nations. 1990. "Step-by-step guide to the estimation of child mortality", *Population Studies*, **107**:1-55.
- United Nations. 1992. *Child mortality since the 1960s: A database for developing countries*. New York: United Nations.
- United Nations Population Division. 2009. *World Population Prospects: The 2008 Revision Population Database*. New York: Department of Economic and Social Affairs.
<http://esa.un.org/wpp/sources/country.aspx>.
- Ward, P. and Zaba, B. 2008. "The effect of HIV on the estimation of child mortality using the children surviving/children ever born technique", *Southern Africa Journal of Demography*, **11** (1): 39-73.
- Zaba, B., Marston, M. and Floyd, S. 2003. "The effect of HIV on child mortality trends in Sub-Saharan Africa " Training workshop on HIV/AIDS and adult mortality in developing countries. New York, 8-13 September 2003

APPENDICES

APPENDIX A Comparison of CEB/CS estimates and corrected IMR and U5MR (Using Hallett estimates for the period 1-5 years before the survey)

Figure A.5.1 Comparison of CEB/CS estimates and corrected IMR and U5MR for Kenya

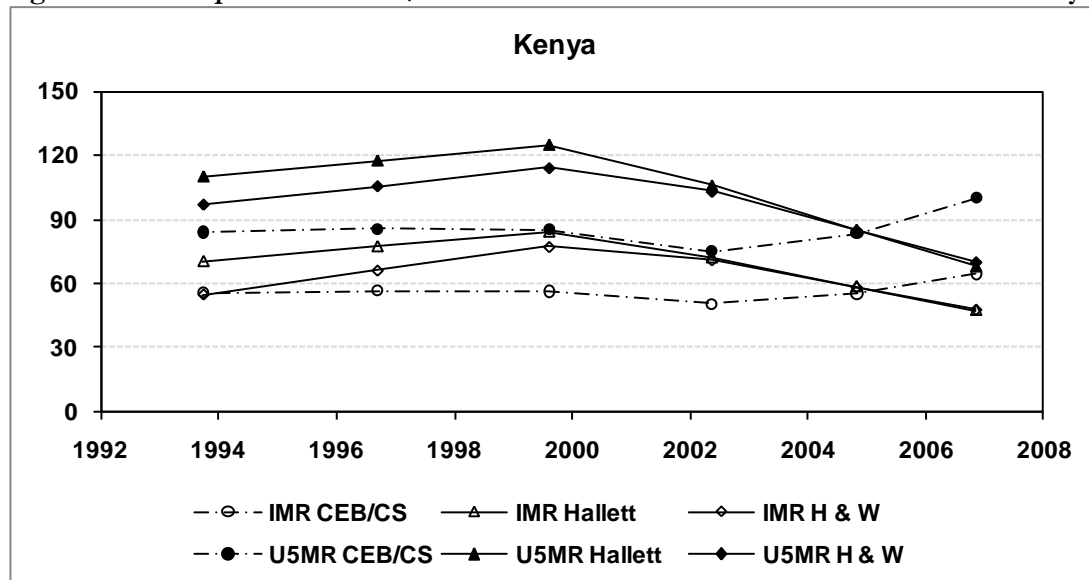


Figure A.5.2 Comparison of CEB/CS estimates and corrected IMR and U5MR for Lesotho

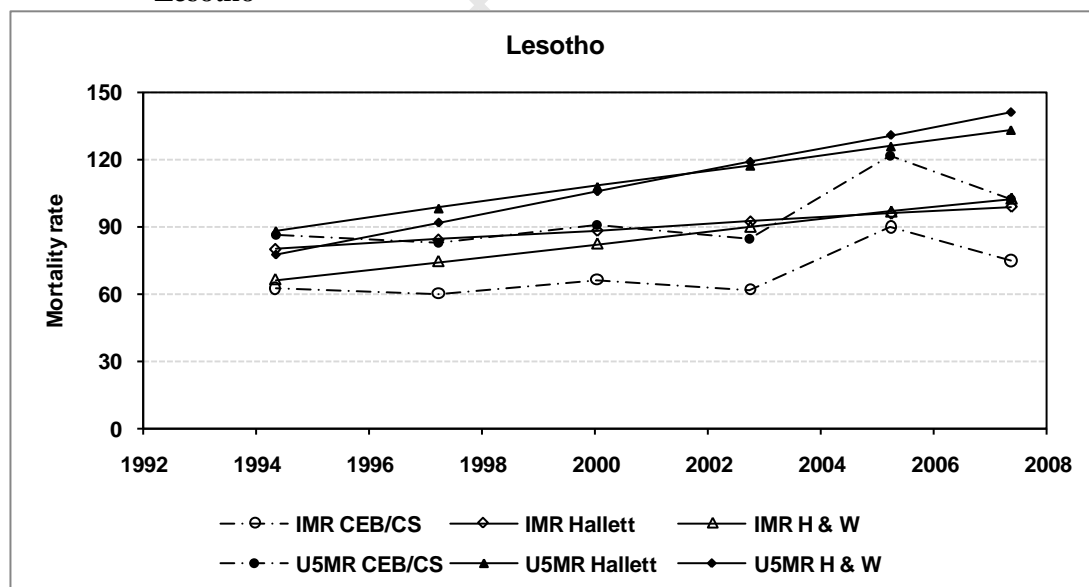


Figure A.3 Comparison of CEB/CS estimates and corrected IMR and U5MR for Malawi

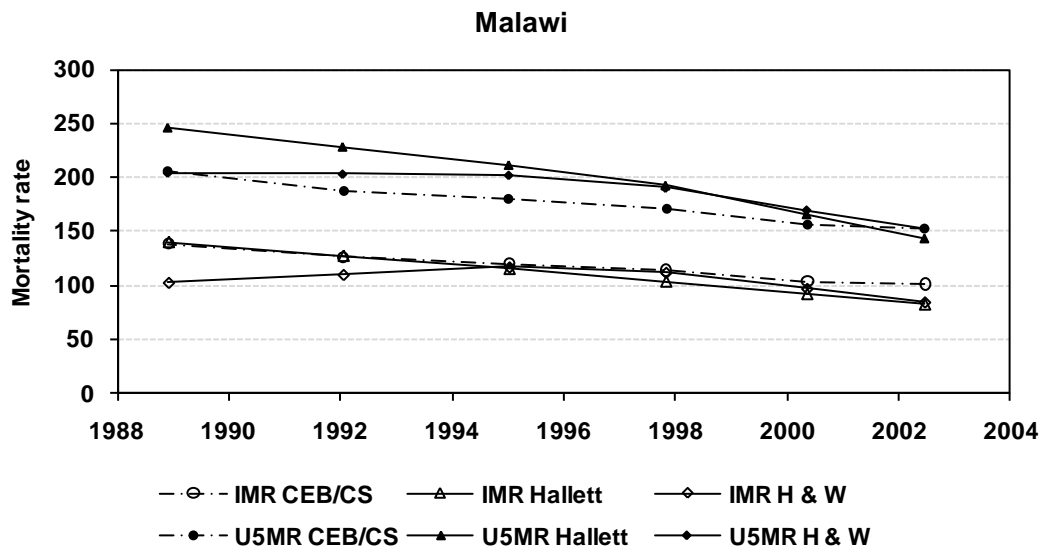


Figure A.4 Comparison of CEB/CS estimates and corrected IMR and U5MR for Zambia

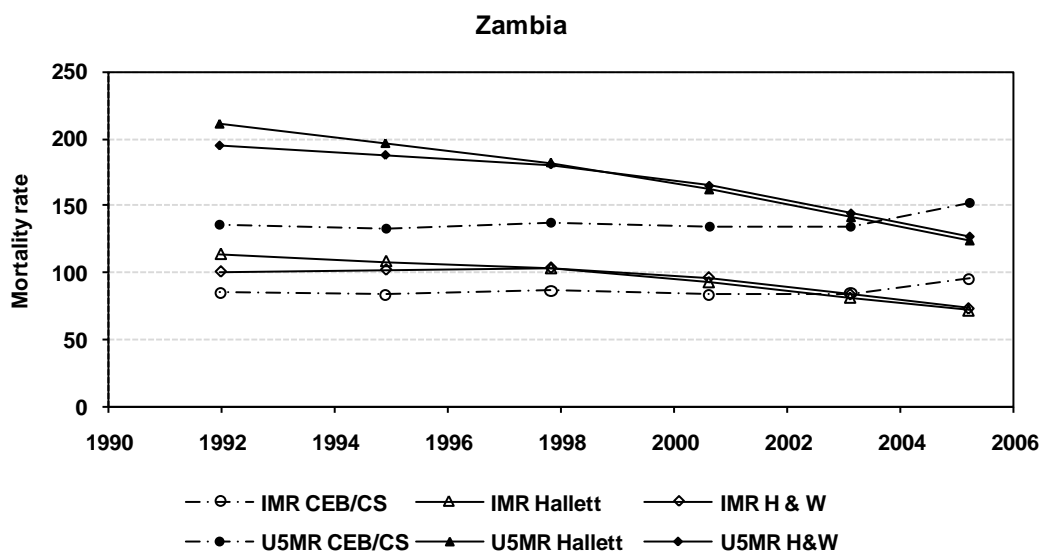
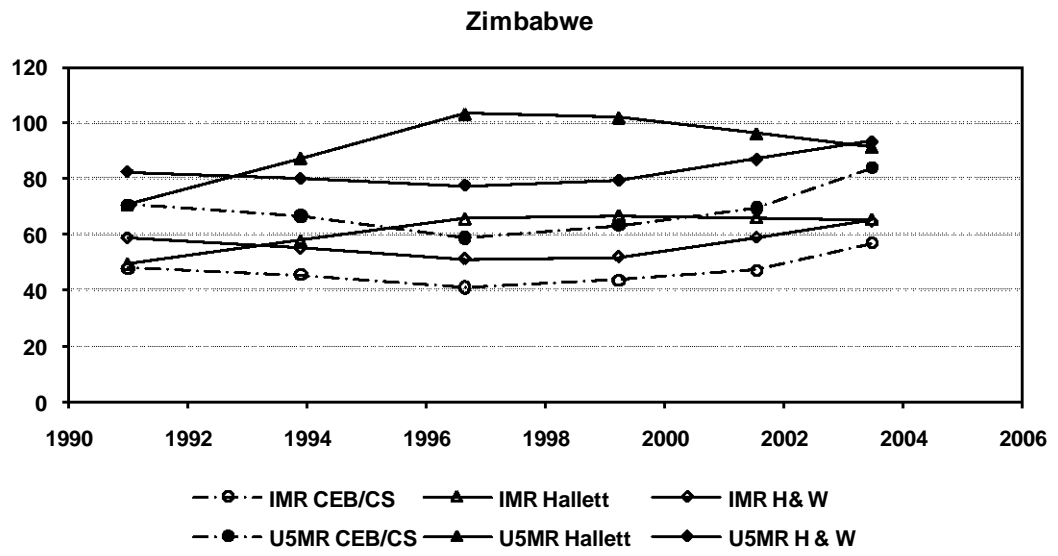


Figure A.5 Comparison of CEB/CS estimates and corrected IMR and U5MR for Zimbabwe



APPENDIX B: Comparison of CEB/CS estimates and corrected IMR and U5MR (using Hallett estimates for the 0-4 year period before the survey)

Figure B.5.3 Comparison of CEB/CS estimates and corrected IMR and U5MR for Kenya

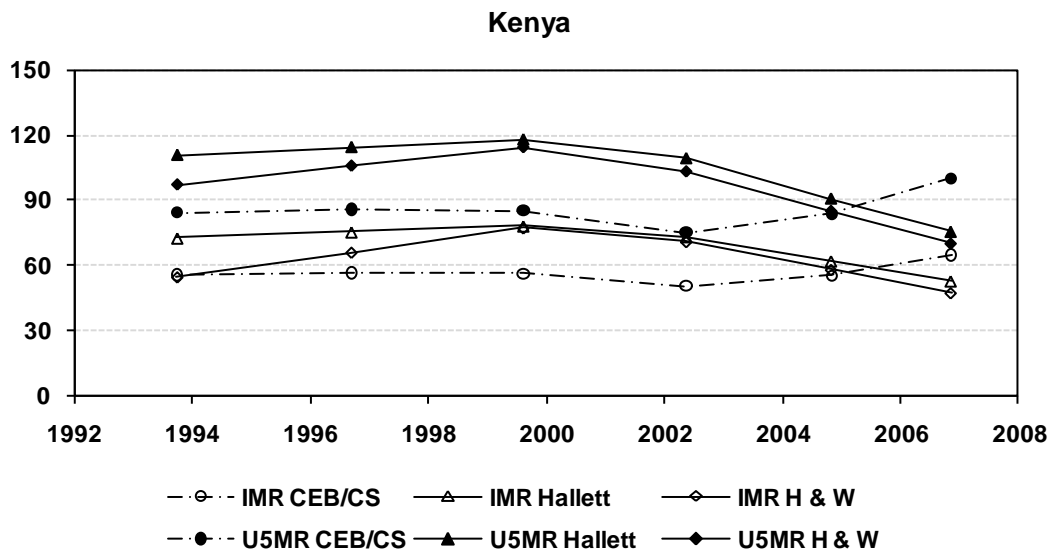


Figure B.5.4 Comparison of CEB/CS estimates and corrected IMR and U5MR for Lesotho

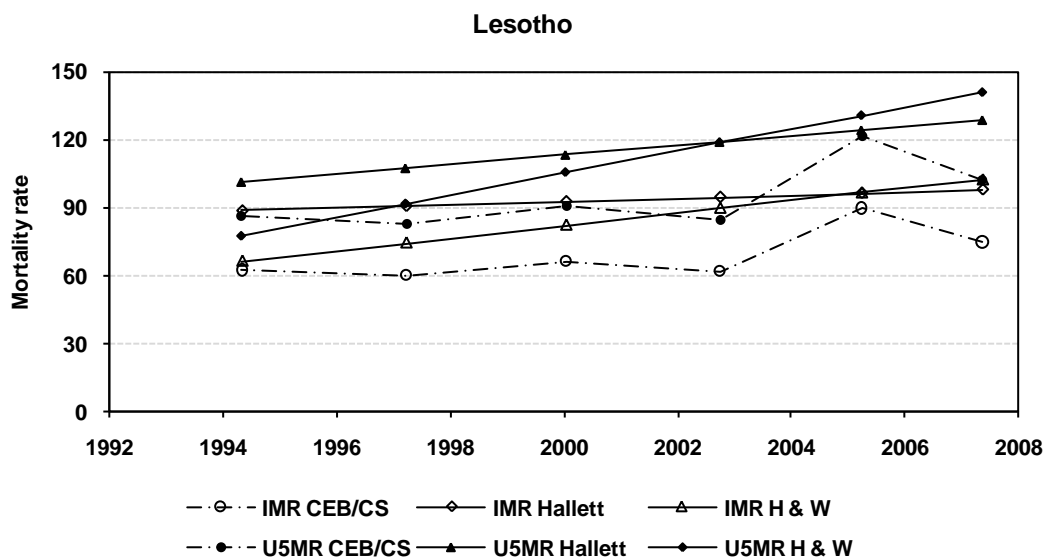


Figure B.3 Comparison of CEB/CS estimates and corrected IMR and U5MR for Malawi

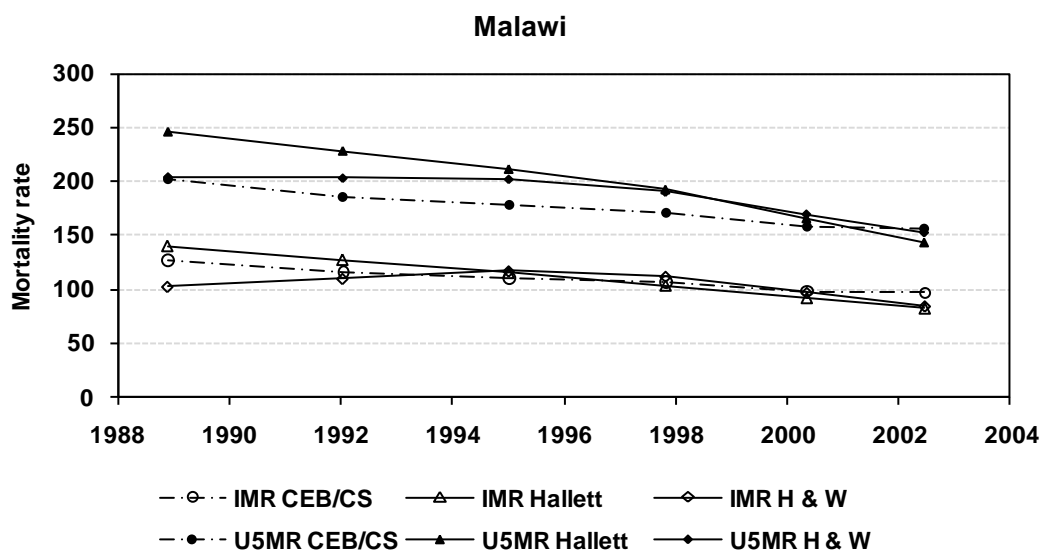


Figure B.4 Comparison of CEB/CS estimates and corrected IMR and U5MR for Zambia

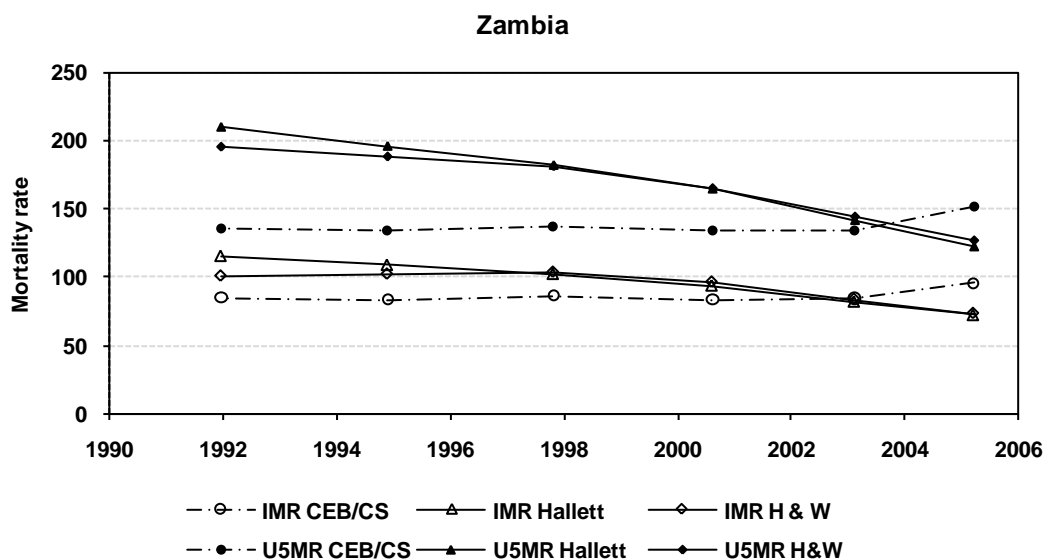
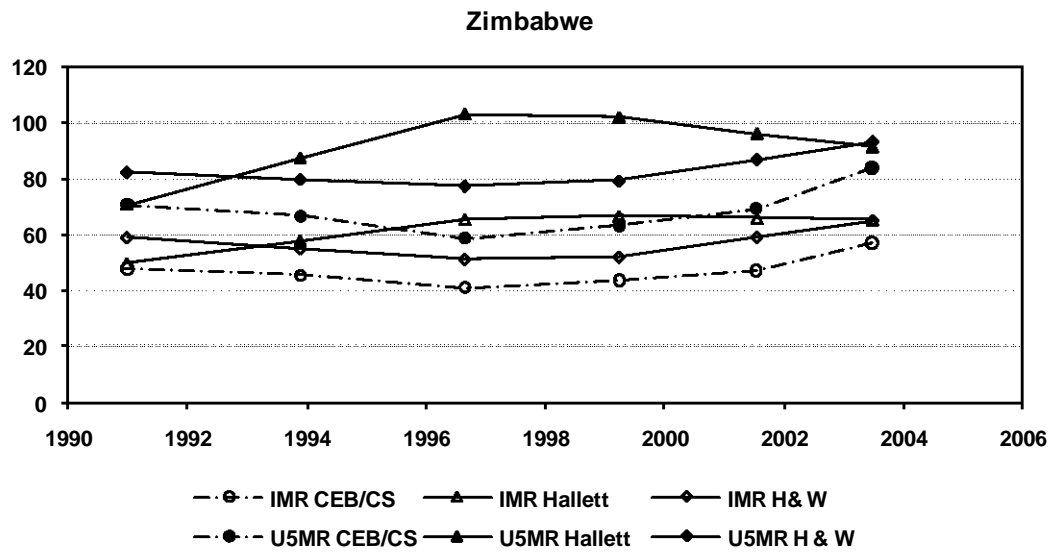


Figure B.5 Comparison of CEB/CS estimates and corrected IMR and U5MR for Zimbabwe



APPENDIX C: Components of bias in infant and under-five mortality estimates derived using the CEB/CS method

Figure C.1 Bias in estimates of IMR and U5MR derived using CEB/CS method for Kenya

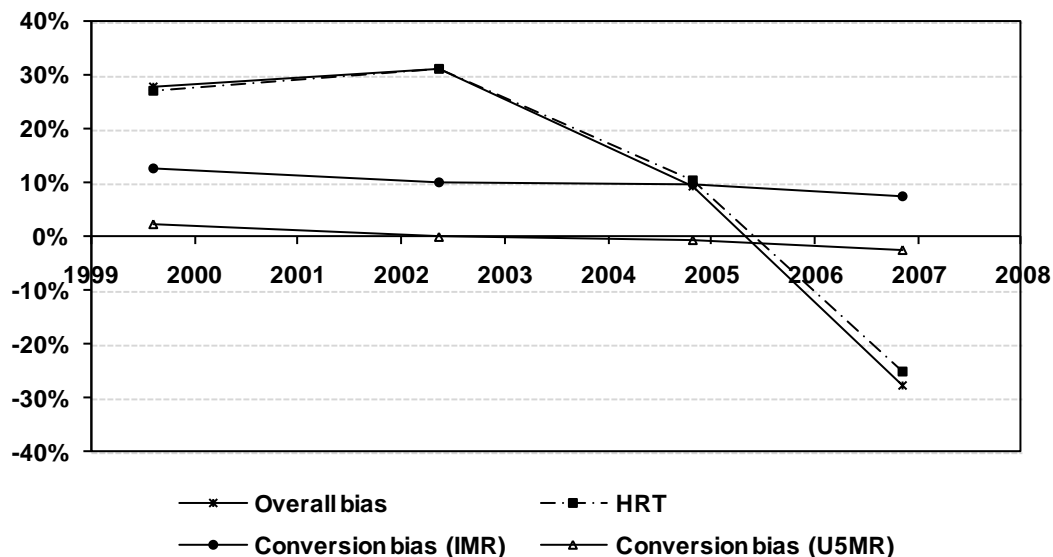


Figure C.2 Bias in estimates of IMR and U5MR derived using CEB/CS method for Lesotho

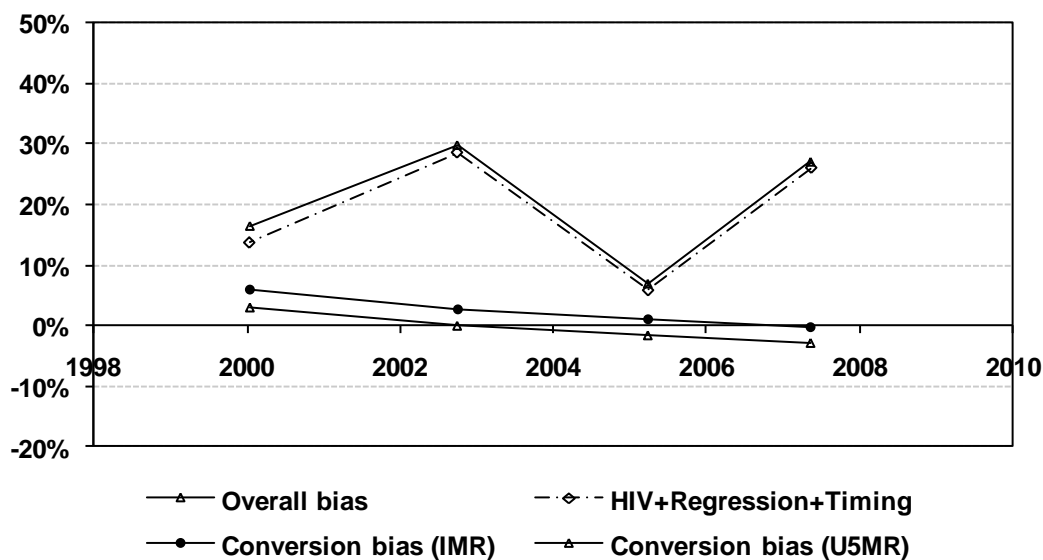


Figure C.3 Bias in estimates of IMR and U5MR derived using CEB/CS method for Malawi

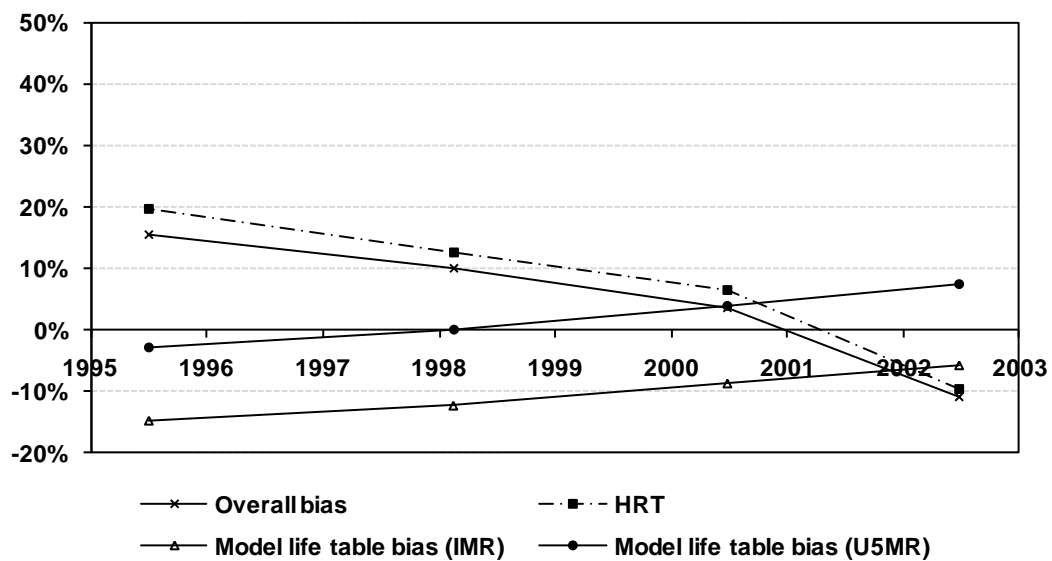


Figure C.4 Bias in estimates of IMR and U5MR derived using CEB/CS method for Namibia

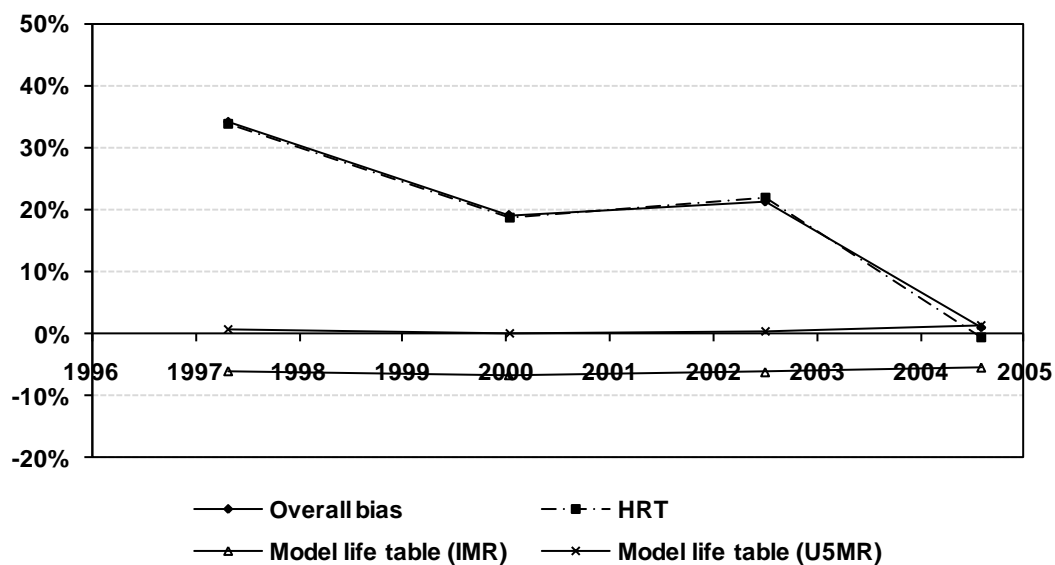


Figure C.5 Bias in estimates of IMR and U5MR derived using CEB/CS method for Zambia

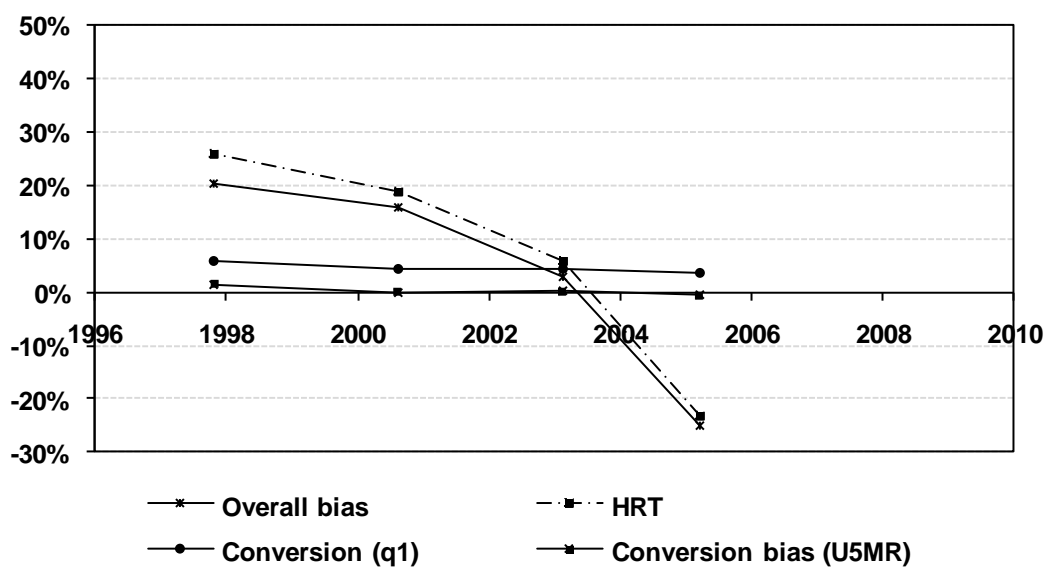


Table C.5.1 Conversion bias in IMR by country and age group of women

Age Group	Country				
	Kenya	Lesotho	Malawi	Namibia	Zambia
20-24	8	0	-6	-5	4
25-29	10	1	-9	-6	4
30-34	10	3	-12	-7	4
35-39	13	6	-15	-6	6

APPENDIX D: Comparison of overall bias from Hallett *et al.* and Hill and Walker methods

Figure D.1 Overall bias from Hallett *et al.* and Hill and Walker methods for Kenya

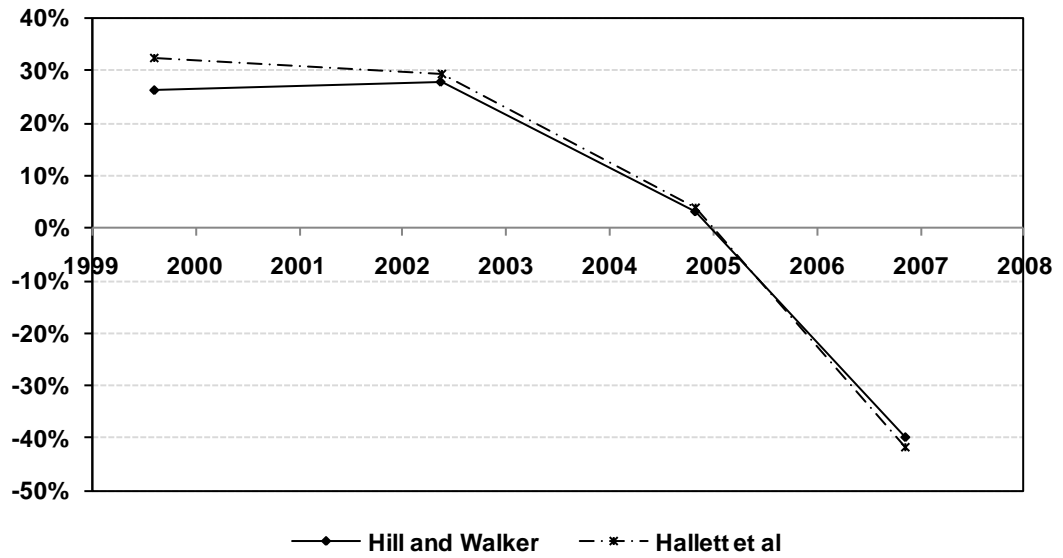


Figure D.2 Overall bias from Hallett *et al.* and Hill and Walker methods for Lesotho

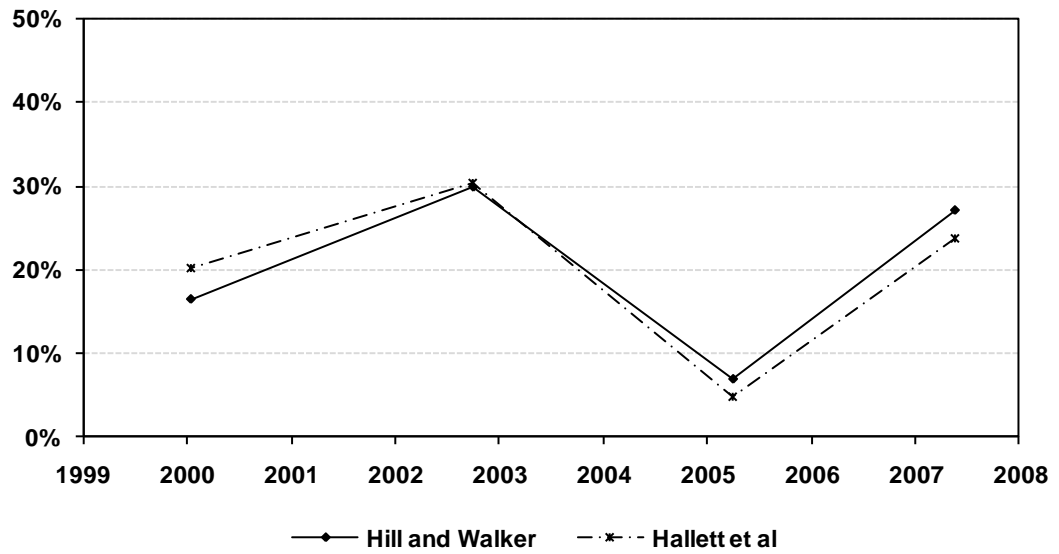


Figure D.3 Overall bias from Hallett *et al.* and Hill and Walker methods for Malawi

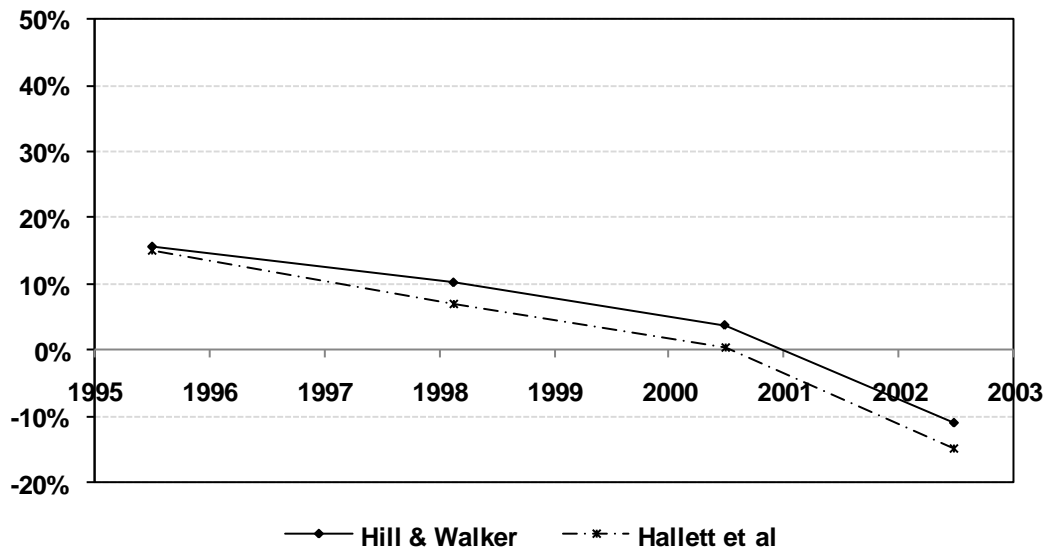


Figure D.4 Overall bias from Hallett *et al.* and Hill and Walker methods for Namibia

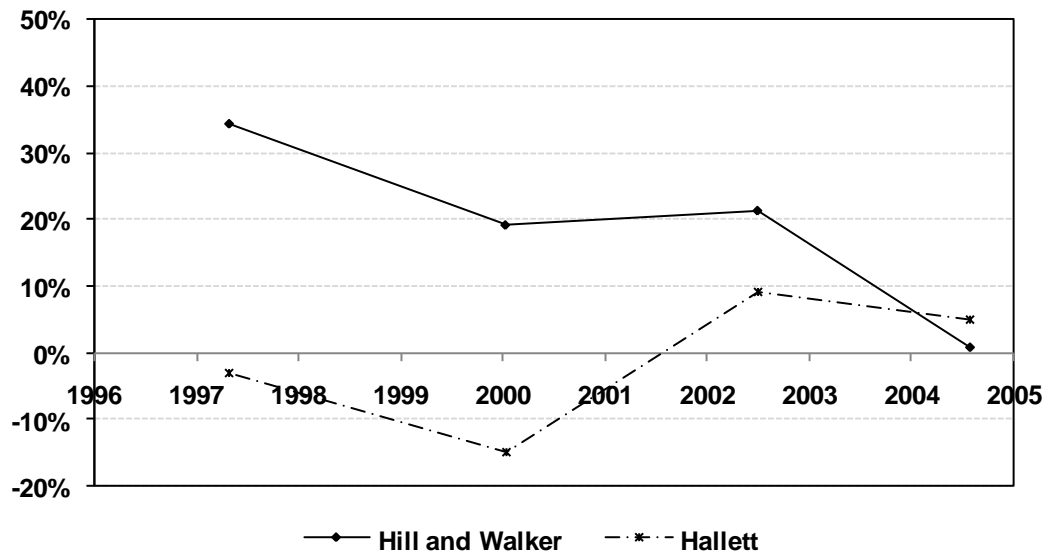


Figure D.5 Overall bias from Hallett *et al.* and Hill and Walker methods for Zambia

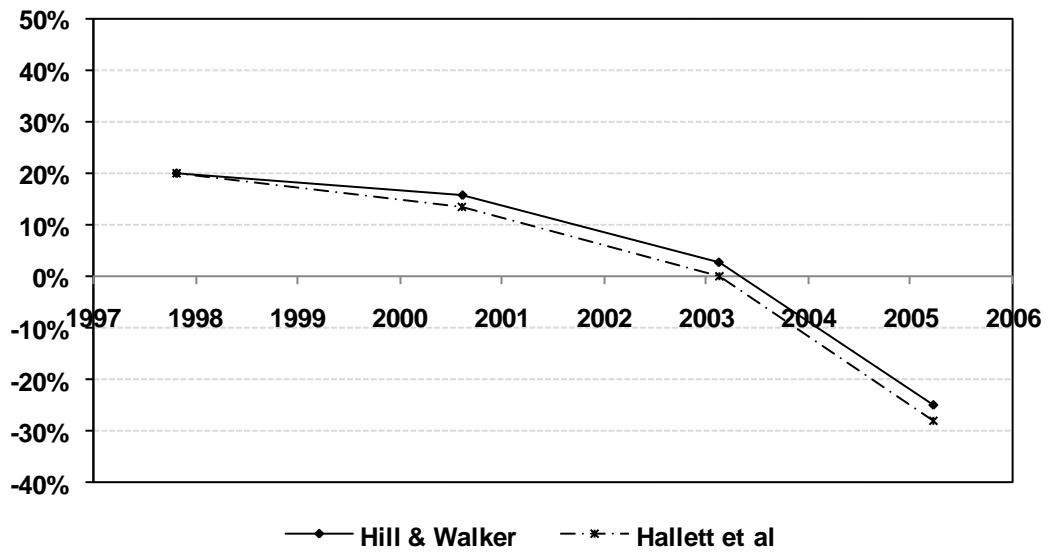
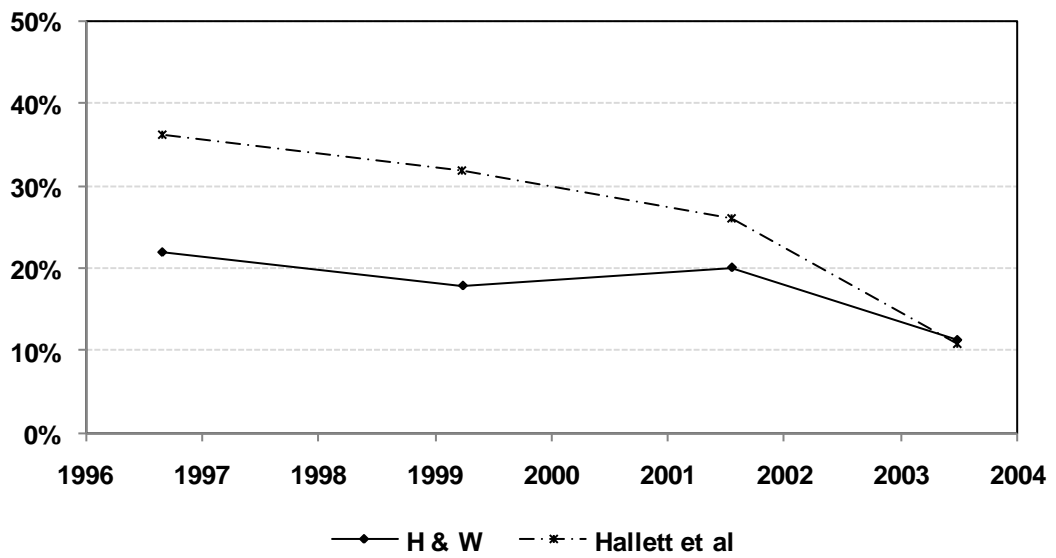


Figure D.6 Overall bias from Hallett *et al.* and Hill and Walker methods for Zimbabwe

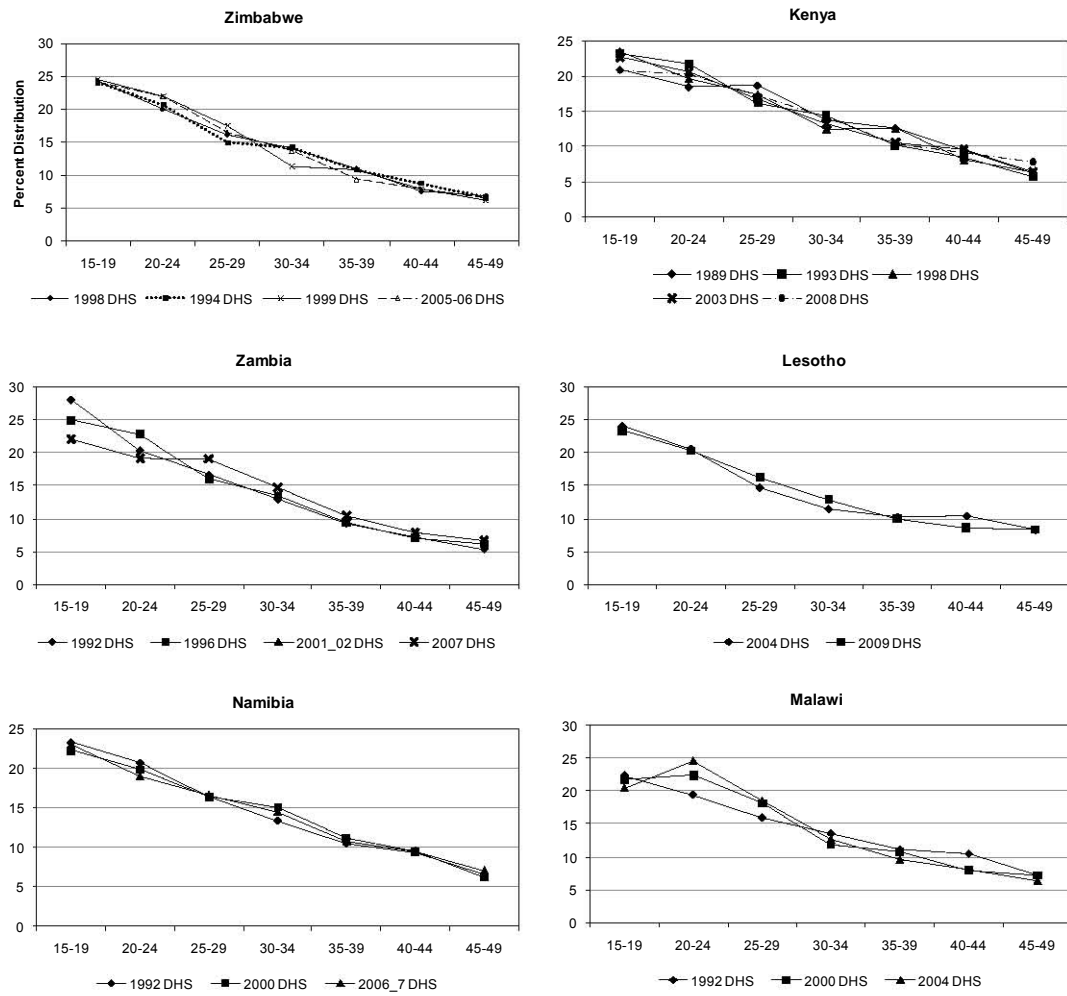


APPENDIX E. DHS Sample sizes by country

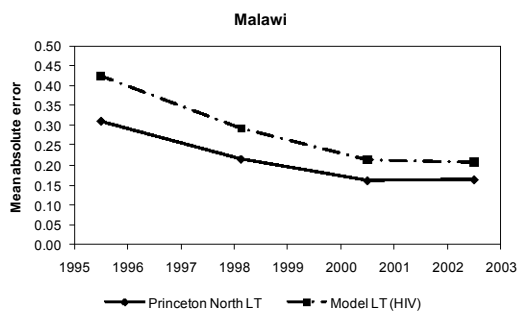
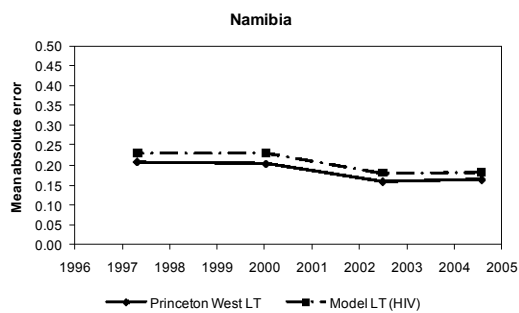
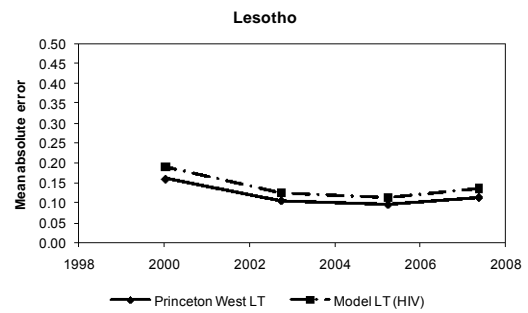
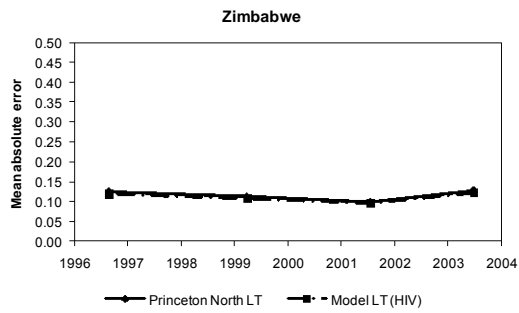
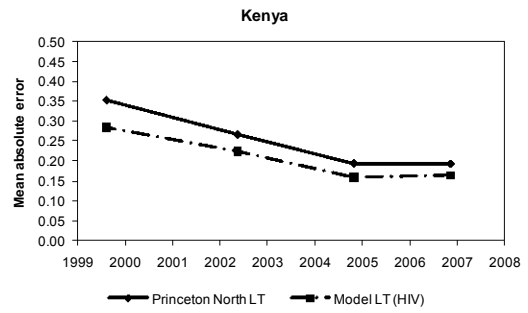
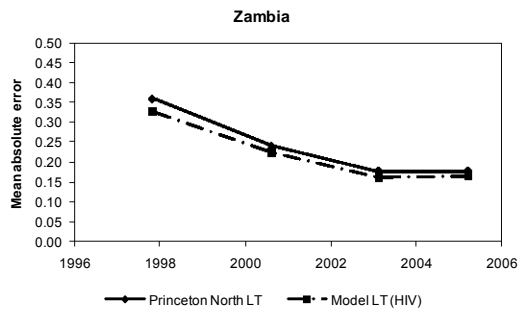
Table E.5.2 DHS samples sizes by country and survey

Country	DHS	Sample Size (No. of Women)
Lesotho	2004	7095
	2009	7624
Kenya	1989	7150
	1993	7540
	1998	7881
	2003	8195
	2008	8444
Malawi	1992	4849
	2000	13220
	2004	11698
Namibia	1992	5421
	2000	6755
	2006-07	9804
Zambia	1992	7060
	1996	8021
	2001-02	7658
	2007	7146
Zimbabwe	1988	4201
	1994	6128
	1999	5907
	2005-06	8907

APPENDIX F: Age distributions of women aged 15-49 by DHS and country



APPENDIX G: Mean absolute error between $q(x)$ values fitted using Princeton and Model Life Tables incorporating HIV



APPENDIX H: Stata program for estimating infant and under-five mortality

```
* To estimate childhood mortality from DHS birth histories
*Creating Child files from mothers' files
cd "c:\data sets"
clear
capture log close
set more off
set mem 300m
log using Zimbabwe_child_2005_06.log, replace
* Create unique identifier for mothers
* By Sorting on cluster, household & line number
use Zimbabwe2005
gen int mumid = _n
label var mumid "unique identifier for mother"
keep v005 v008 v013 v025 v106 v201 mumid bidx* b2* b3* b4* b5* b6* b7* b8*
* Renaming variables
rename v005 weight
rename v008 interviewcmc
rename v013 mumagrp
rename v025 mumurbru
rename v106 mumeduc
rename v201 mumceb
renpfix bidx_0 kidid
renpfix b2_0 kidyob
renpfix b3_0 kidcmc
renpfix b4_0 kidsex
renpfix b5_0 kidalive
renpfix b6_0 kiddthcode
renpfix b7_0 kidmondth
renpfix b8_0 kidaglbday
renpfix bidx_ kidid
renpfix b2_ kidyob
renpfix b3_ kidcmc
renpfix b4_ kidsex
renpfix b5_ kidalive
renpfix b6_ kiddthcode
renpfix b7_ kidmondth
renpfix b8_ kidaglbday
* Dropping records for mothers who never gave birth
drop if mumceb==0
browse mum* kidid1 kidyob1 kidsex1 kidid2 kidyob2 kidsex2 kidid3 kidyob3 kidsex3
kidid4 kidyob4 kidsex4 kidid5 kidyob5 kidsex5 kidid6 kidyob6 kidsex6 kidid7 kidyob7
kidsex7 kidid8 kidyob8 kidsex8 kidid9 kidyob9 kidsex9 kidid10 kidyob10 kidsex10
* Reshape the file to get one record per child, Drop unused records
reshape long kidid kidyob kidcmc kidsex kidalive kiddthcode kidmondth kidaglbday,
i(mumid) j(jay)
browse mum* jay kidid kidyob kidsex
drop if kidid==.
drop jay
save child, replace
```

```

*Program for calculating death
use "c:\Data sets\child.dta"
gen agegr_1 = 0
gen agegr_2 = 1
gen agegr_3 = 3
gen agegr_4 = 6
gen agegr_5 = 12
gen agegr_6 = 24
gen agegr_7 = 36
gen agegr_8 = 48
gen agegr_9 = 60

*set width of each period for analysis (in months)
*program works for minimum period of 12 months
gen period = 60
*set number of periods for analysis
gen maxper = 6
*set upper and lower limits for date of analysis period
gen uplim = interviewcmc - 1
gen lowlim = interviewcmc - (maxper*period)- 1
*select only dead children born in the periods of analysis
gen xproc = 0
replace xproc = 1 if lowlim <=kidcmc & kidcmc <= uplim & kidalive == 0
keep if xproc ==1
*Assign deaths to age groups
gen j = 0
*Age at deaths = 0 months
replace j = 1 if agegr_1 <=kidmondth & kidmondth < agegr_2
*Age at death = 1-2 months
replace j = 2 if agegr_2 <=kidmondth & kidmondth < agegr_3
*Age at death = 4-5 months
replace j = 3 if agegr_3 <=kidmondth & kidmondth < agegr_4
*Age at death = 6-11 months
replace j = 4 if agegr_4 <=kidmondth & kidmondth < agegr_5
*Age at death = 12-23 months
replace j = 5 if agegr_5 <=kidmondth & kidmondth < agegr_6
*Age at death = 24-35 months
replace j = 6 if agegr_6 <=kidmondth & kidmondth < agegr_7
*Age at death = 36-47 months
replace j = 7 if agegr_7 <=kidmondth & kidmondth < agegr_8
*Age at death = 48-59 months
replace j = 8 if agegr_8 <=kidmondth & kidmondth < agegr_9
gen agedth = j - 1
*Select children who died under age 5
keep if j != 0
*Determine period of birth and death
gen perborn = int((interviewcmc-1-kidcmc)/period)
*Calculate lower bound for the date of the period in which the child was born (limlow)
gen limlow = interviewcmc - (perborn+1)*period
*Calculate earliest date death could occur in age group j

```

```

gen agei = 0
*i.e month of birth
replace agei = kidcmc+agegr_1 if j==1
*i.e month of birth + 1 month
replace agei = kidcmc+agegr_2 if j==2
*i.e month of birth + 3 months
replace agei = kidcmc+agegr_3 if j==3
*i.e month of birth + 6 months
replace agei = kidcmc+agegr_4 if j==4
*i.e month of birth + 12 months
replace agei = kidcmc+agegr_5 if j==5
*i.e month of birth + 24 months
replace agei = kidcmc+agegr_6 if j==6
*i.e month of birth + 36 months
replace agei = kidcmc+agegr_7 if j==7
*i.e month of birth + 48 months
replace agei = kidcmc+agegr_8 if j==8
*Calculate date of start of next age group
*(i.e. upper bound on date of death in age group j)
gen nxtage = 0
*i.e month of birth + 1 month
replace nxtage = kidcmc+agegr_2 if j==1
*i.e month of birth + 3 months
replace nxtage = kidcmc+agegr_3 if j==2
*i.e month of birth + 6 months
replace nxtage = kidcmc+agegr_4 if j==3
*i.e month of birth + 12 months
replace nxtage = kidcmc+agegr_5 if j==4
*i.e month of birth + 24 months
replace nxtage = kidcmc+agegr_6 if j==5
*i.e month of birth + 36 months
replace nxtage = kidcmc+agegr_7 if j==6
*i.e month of birth + 48 months
replace nxtage = kidcmc+agegr_8 if j==7
*i.e month of birth + 60 months
replace nxtage = kidcmc+agegr_9 if j==8
*Calculate upper bound for the date of the period in which the child was born (limupp)
gen limupp = limlow+period
gen n = 1
*Number of periods in which death could occur
gen iter = 0
*Death could occur in same period of birth
replace iter = 1 if limlow<= kidcmc &nxtage <= limupp
*Death could occur in period of birth or in the next period
replace iter = 2 if agei < limupp & limupp <= nxtage
*Death occurs in period after birth
replace iter = 1 if kidcmc< limupp & limupp <= agei
*Set perborn to period of death, i.e. next period
replace perborn = perborn - 1 if kidcmc< limupp & limupp <= agei
*All deaths to children born in the most recent period must occur in the most recent
period

```

```

replace iter = 1 if perborn == 0
replace n = n/iter if iter !=0
*Colper defines columns for table = time periods
gen colper = perborn
*Weight the data. Division by 10 used instead of 1000000 to effectively allow more
decimal places in the tabulation.
*Deaths that could have occurred in either of the two time periods are assigned 1/2 to
each period (n)
gen rweight = n*weight/10
*Tabulate deaths that occurred to children born in the last 5 periods by age at death and
period
gen xtabs = 0
replace xtabs = 1 if iter != 0 & 0 <= colper & colper < 5
* Output for death
save "C:\death.dta", replace
keep if xtabs ==1
tabulate agedth colper [iweight=rweight]
clear
*Retabulate deaths that could have occurred in the next period, in that period
use "c:\death.dta"
replace colper = colper-1 if iter == 2
gen xtabs1 = 0
replace xtabs1 = 1 if iter == 2 & 0 <=colper & colper < 5
keep if xtabs1 == 1
tabulate agedth colper [iweight=rweight]
clear
*Combine the two system files of death into one file
insheet using "C:\combine.csv", comma
tabulate dthage clper [fweight = dths]
clear
insheet using "C:\death.csv", comma
sort dthage clper
*Save deaths as STATA data file
save "C:\deaths.dta",replace
clear
use "C:\child.dta"
*Set width of each period for analysis. Minimum = 12 months
gen period=60
*Set number of periods for analysis
gen maxper = 6
*Define lower limits of age categories for calculating probabilities
gen agegr_1= 0
gen agegr_2= 1
gen agegr_3= 3
gen agegr_4= 6
gen agegr_5= 12
gen agegr_6= 24
gen agegr_7=36
gen agegr_8=48
gen agegr_9=60
*Set upper and lower limits for date of analysis period

```

```

gen limupp= interviewcmc-1
gen limlow = interviewcmc-(maxper*period)-1
*Select children born in the analysis period
gen xproc = 0
replace xproc=1 if limlow<=kidcmc & kidcmc<=limupp
keep if xproc==1
*Set months = number of months child lived
gen months=0
replace months=kidmondth if kidalive==0
replace months=interviewcmc-kidcmc if kidalive==1
*Calculate period of birth
gen perborn=int((interviewcmc-1-kidcmc)/period)
save "C:\chn.dta", replace

*Tabulate exposure in the first age group (0 months) by period
gen ageexp=0
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc
gen nxtage=kidcmc+agegr_2

do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear

*Tabulate exposure in the second age group (1-2 months) by period
use "C:\chn.dta"
gen ageexp=1
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_2
gen nxtage=kidcmc+agegr_3
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear

*Tabulate exposure in the third age group (3-5 months) by period
use "C:\chn.dta"
gen ageexp=2
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_3
gen nxtage=kidcmc+agegr_4
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear

*Tabulate exposure in the fourth age group (6-11 months) by period
use "C:\chn.dta"
gen ageexp=3
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_4
gen nxtage=kidcmc+agegr_5
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear

```

```

*Tabulate exposure in the fifth age group (12-23 months) by period
use "C:\chn.dta"
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen ageexp=4
gen agei=kidcmc+agegr_5
gen nxtage=kidcmc+agegr_6
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear
*Tabulate exposure in the sixth age group (24-35 Months) by period
use "C:\chn.dta"
gen ageexp=5
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_6
gen nxtage = kidcmc+agegr_7
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear
*Tabulate exposure in the seventh age group (36-47 months) by period
use "C:\chn.dta"
gen ageexp=6
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_7
gen nxtage=kidcmc+agegr_8
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear
*Tabulate exposure in the eighth age group (48-59 months) by period
use "C:\chn.dta"
gen ageexp=7
*Set agei to CMC for start of age group and next age to CMC for start of next age group
gen agei=kidcmc+agegr_8
gen nxtage=kidcmc+agegr_9
do "C:\do\mort2.do"
do "C:\do\mort3.do"
clear
insheet using "C:\expo1.csv", comma
tabulate expage col [fweight = exp]
clear
insheet using "C:\exposure.csv", comma
save "C:\exposure.dta", replace
clear
*Define the period of birth of the child and the number of iterations
*Select children exposed for at least part of the age group
*i.e. children who enter the age group
keep if agei <= kidcmc+months
*Calculate lower bound for the date of the period in which the child was born
gen lowlim = interviewcmc-((perborn+1)*period)
*Calculate upper bound for the date of the period in which the child was born
gen upplim = lowlim+period
*Determine number of periods in which exposure occurred in the age group (iter)

```

```

gen iter = 0
replace perborn = perborn-1 if upplim <= agei
replace iter = 1 if upplim <= agei
gen n = 0
replace n = 1 if upplim<= agei
replace lowlim = lowlim+ period if upplim <= agei
replace upplim = lowlim+ period if upplim <= agei
*All exposure occur in months of birth
replace iter = 1 if nxtage < upplim
replace n = 1 if nxtage < upplim
*Exposure occurs in period of birth and in the next period
replace iter = 2 if agei < upplim & upplim <= nxtage
replace n = 0.5 if agei < upplim & upplim <= nxtage
replace perborn = perborn-1 if upplim <= agei
replace iter = 1 if upplim <= agei
replace iter = 1 if agei < upplim & upplim <= nxtage & perborn == 0
*Colper defines columns for tabulation = time periods
gen colper = perborn
*Weight data
*Division by 100 used to all more decimal places in the tabulation.
*Exposure that occurs over two time periods is assigned 1/2 to each period (n)
gen rweight = n*weight/100
*Select 5 periods for tabulation
gen xproc1 = 0
replace xproc1 = 1 if 0 <= colper & colper < 5
*Tabulate the first part of the exposure
save "C:\mort2.dta", replace
keep if xproc1 == 1
tabulate ageexp colper [iweight=rweight]
clear
*Prepare for the second part of the tabulation
use "C:\mort2.dta"
gen colper1 = colper - 1
gen xproc2 = 0
replace xproc2 = 1 if 0 <= colper1 & colper1 <= 4 & iter == 2
*Tabulate the second part of the exposure
keep if xproc2 == 1
tabulate ageexp colper1 [iweight=rweight]
clear
use "C:\deaths.dta", clear
gen id = _n
order id
sort id
save "C:\Tdeaths.dta", replace
clear
use "C:\exposure.dta", clear
gen id = _n
drop ageexpos
order id
sort id
save "C:\sexposure.dta", replace

```

```

clear
use "C:\Tdeaths.dta", clear
merge id using "C:\sexposure.dta"
rename dthage agemnth
*rename clper colper
save "C:\dthsexp.dta", replace
label data "This file contains exposure and deaths for Malawi 2004"
label variable agemnth " Age in months - Deaths"
label variable colper " Periods of analysis of 5 years"
label variable exposure " exposure to the risk of dying"
label define agemnth1 0 " 0 " 1 "1 - 2" 2 "3 - 5" 3 "6 - 11" 4 "12 - 23" 5 "24 - 35" 6 "36 -
47" 7 "48 - 59"
label values agemnth agemnth1
label define colper1 0 " 0 - 4 " 1 "5 - 9" 2 "10 - 14" 3 "15 - 19" 4 "20 - 24"
label values colper colper1
*Tabulate deaths by age group and period
gen rweight= dths/1000
tab agemnth colper [iweight=rweight]
*Tabulate exposure by age group and period
label variable agemnth " Age in months - exposure"
gen rweight1 = exposure/1000
tab agemnth colper [iweight=rweight1]
*Tabulate probabilities by age group and period
gen agepro=agemnth
label variable agepro " Age in months - probabilities(times 1000000)"
label define agepro1 0 " 0 " 1 "1 - 2" 2 "3 - 5" 3 "6 - 11" 4 "12 - 23" 5 "24 - 35" 6 "36 -
47" 7 "48 - 59"
label values agepro agepro1
gen rweight2 = dths*1000000/exposure
tab agepro colper [iweight=rweight2]
clear
insheet using "C:\Probs.csv"
gen probdc = 0
gen mn = 0
gen infant =0
*Mortality indicates each mortality rate e.g. 0 = neonatal, 2 = infant, 3 = child, and is set
at the last age group for each rate
gen mortrate = 99
replace mortrate = 0 if ageprobs ==0
replace mortrate = 2 if ageprobs ==3
replace mortrate = 3 if ageprobs ==7
*Calculate cumulative probability of surviving to the end of each age group (nP0)
replace probdc = 10000000-probs if mortrate==0
gen id =_n
sort id
tsset id
replace probdc = ((L.probdc)*(10000000-probs))/10000000 if mortrate!=0
*For age groups 4 (12-23) upwards, calculate probability of surviving from age 12
months
*to the end of each age group (nP12) to get the child mortality rate. Therefore reset
cumulative probability

```

```

*for age group 12-23 months to start cumulating from that age group on
replace probdc = 10000000-probs if ageprobs ==4
replace probdc = ((L.probdc)*(10000000-probs))/10000000 if mortrate!=0 &
ageprobs!=4
*mn = probability of surviving 1st month, infant = probability of surviving
replace mn = probdc if mortrate==0
replace mn = L.mn if mortrate !=0
replace infant = probdc if mortrate ==2
replace infant = L.infant if mortrate!=2
*Compute probability of death (nqx) from probability of surviving
gen rate = 10000000-probdc
*Tabulate mortality rates for neonatal, infant, and child mortality by period and save
save "C:\neonat.dta", replace
keep if mortrate!=99
tabulate colper mortrate [iweight=rate]
clear
*Calculate postneonatal mortality rate (mortrate = 1) and under-five mortality rate
(mortrate = 4)
use "C:\neonat.dta"
gen mortrate1=99
replace mortrate1 = 1 if ageprobs ==3
replace mortrate1 = 4 if ageprobs==7
replace rate = (10000000-probdc)-(10000000-mn) if mortrate1==1
replace rate = 10000000-((probdc*infant)/10000000) if mortrate1==4
*Tabulate PNMR and U5MR by period
keep if mortrate1!=99
tabulate colper mortrate1 [iweight=rate]
clear
insheet using "C:\rates.csv"
order rates
label variable mortrate " Mortality Rates"
label define mortrate1 0 " Neonatal (NM)" 1 "Post-neo (PNM)" 2 "Infant (1q0)" 3
"Child (4q1)" 4 "Under-five (5q0)"
label values mortrate mortrate1
label variable colper " Periods of analysis of 5 years"
label define colper1 0 " 0 - 4" 1 "5 - 9" 2 "10 - 14" 3 "15 - 19" 4 "20 - 24"
label values colper colper1
gen rweight6 = rate/10000
/*Infant and Child Mortality DHS Estimates for ZDHS2005*/
tabulate colper mortrate [iweight=rweight6]

```