Quantifying Selection Bias from Birth History Estimates of Child Mortality

Matthew Coates

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2016

Committee:

Haidong Wang

Emmanuela Gakidou

Laura Dwyer-Lindgren

Program Authorized to Offer Degree:

Global Health

University of Washington

Abstract

Quantifying Selection Bias from Birth History Estimates of Child Mortality
Matthew Coates

Chair of the Supervisory Committee:
Haidong Wang, Associate Professor
Global Health

Introduction Child mortality rates have long been used as an indicator of progress in global health and development. Their measurement in developing settings is often based on birth history data provided by mothers in surveys; however, these measures miss the mortality risk in orphans due to selection bias in only sampling mothers who are alive. Methods exist to address this bias using HIV projection models, but comprehensive estimates of the bias from nationally representative survey data do not exist.

Methods I have developed an approach to estimating this selection bias, using sibling survival methods to estimate the ratio of orphan to non-orphan mortality and household survey data to estimate the prevalence of maternal orphans. In particular, I have focused on 48 Demographic and Health Surveys (DHS) in Sub-Saharan Africa and applied established methods to estimate a correction factor.

Results This method has found significant, though often small bias in child mortality estimates using complete birth histories in 36 out of 48 surveys. The largest bias was found in surveys from Namibia and Zimbabwe with around a 10% increase in under-5 mortality estimates when accounting for the bias.

Discussion The method shows mixed consistency with the existing methods, and several challenges exist to producing reliable estimates. Sibling survival sample sizes for the groups required by this analysis are often insufficient, and biases inherent to sibling survival methods in estimating child mortality need to be further addressed. Extensions of this method offer the potential for using nationally representative survey data to create geographically comprehensive adjustments for complete birth history estimates of child mortality due to selection bias.

Table of Contents

Introduction

Child and infant mortality rates have been suggested by the global health and development communities as indicators to track progress and inequality across countries for some time.[1,2] In recent years, the Millennium Development Goals (MDGs) have brought extra attention to under-five mortality, with MDG 4 setting a goal for two-thirds reduction in under-five mortality rates between 1990 and 2015.[3] While marked progress has been made, with varying sources reporting that almost 60 countries have met this goal, disparities remain.[4,5] Twenty-three countries would need to reduce under-5 mortality by more than 50 deaths per 1000 live births in order to meet the new Sustainable Development Goal (SDG) of 25 deaths per 1000 births.[5,6]

Although under-5 mortality is a heavily-used indicator, high levels of confidence in estimates are not always readily attainable. In fact, when accounting for differences in which specific countries met the MDG 4 goal as opposed to looking only at the total number of about 60 countries that were estimated to have met the goal, only 42 countries succeeded according to both the Global Burden of Disease (GBD) Study 2015 and the United Nations Inter-agency Group for Child Mortality Estimation (IGME) 2015 estimates.[5] Estimates can diverge for several reasons. Often, inclusion of different surveys and different techniques of obtaining estimates from these surveys can yield differing results, and choices about which sources are the most reliable can also affect the final estimates generated by models from the respective groups.

Complete Birth History Approach

In developing countries, estimates of under-5 mortality often rely on survey methods using demographic techniques. So-called *direct* and *indirect* methods can be used, so named because direct estimates can be calculated using numerators and denominators directly from data, while indirect methods rely on assumptions or other data sources to approximate a direct measure. Full or "complete" birth history (CBH) modules that allow for direct calculation of child mortality estimates are often included in large demographic surveys like the Demographic and Health Survey (DHS) or Multiple Indicator Cluster Survey (MICS) series. These questionnaires ask for the date of birth and, if applicable, age at death of each child a sampled mother has had. CBH estimates tend to be used as a standard in countries without vital registration data, but they can be subject to biases. One bias of concern is the compositional bias, with earlier cohorts of children being born to younger women, by nature of the survey being conducted years into the future and fewer older women being sampled. This bias can be addressed in part by using only the 15 years of data prior to the survey.[7] Another source of bias is misreporting– omission of children, incorrect recall of dates, or rounding of children's ages. An additional concern is birth transference, or the movement of children five years of age to six years in order to avoid follow-up questions.[8] Lastly, the samples of these surveys are subject to selection bias. Only women who are alive can be surveyed, thus, mortality among maternal orphans is excluded from CBH estimates of child mortality.[9] While this only creates bias if orphan and non-orphan mortality rates are different, there is both reason to believe this might be the case and evidence to support that assumption.[10–12]

Selection Bias Due to Sampling Alive Mothers

While CBH estimates of under-5 mortality have always been subject to this selection bias, the impact of HIV/AIDS on under-5 mortality has brought more attention to this issue.[13] Ultimately, whether this bias meaningfully affects estimates depends on both the relative mortality rates between orphans and non-

orphans, as well as the prevalence of orphans. Comprehensive estimates of orphan prevalence exist from both surveys and projection models like the UNAIDS Spectrum model.[14–16] Comprehensive estimation of orphan versus non-orphan mortality by country or over time is less widely available, with the most commonly used adjustment method only accounting for orphanhood due to HIV, and assuming rates of vertical transmission and survival from cohort studies and model life tables.[17] To my knowledge, no comprehensive sets of estimates of orphan to non-orphan mortality have been estimated using available nationally representative survey data.

Differentials in orphan and non-orphan under-5 mortality, however, have been documented in various studies.[10–12,18] For example, a cohort study using health and demographic surveillance system data found the probability of survival to age ten among orphans to be less than a third of the probability of survival among non-orphans.[10] Another cohort study in Uganda found the hazard of under-5 mortality to be 3.8 times higher among orphans than non-orphans, similar to the risk estimate of having a mother with HIV (hazard ratio of 3.2).[12] Other studies have been mixed, with some suggesting that a mother's death only poses additional risk if it was a death due to HIV.[18] Hallett et al. (2010) used empirical data and several assumptions to estimate not only the mortality in children who lost mothers to AIDS, but also bias correction factors for DHS estimates of under-5 mortality from complete birth histories using a simulation model.[11] Several proposals and assessments have been made of corrections to both direct and indirect estimation of under-5 mortality in settings with high AIDS mortality.[11,17,19–22]

Two of the large groups creating comprehensive global estimates of child mortality handle this selection bias differently. The GBD studies have made no correction for orphan mortality not captured by CBH methods. Rajaratnam et al. (2010) reference concern that HIV mortality may be higher in children with higher mortality,[23] but conclude on the basis of a subanalysis that the effect of the bias would be small and adjusting would be made difficult by confounding.[24] Although this subanalysis found relatively large differences in child mortality in women living with and without HIV across a range of DHS surveys, the estimated effect of the bias on the final estimates ranged from an underestimation of 10 deaths per 1,000 births to an overestimation of six deaths per 1,000 births. This justification, however, ignores the children who would not have been included in the recall of the survey because their mothers had deceased. While there is certainly bias in ignoring this subpopulation, if the relative mortality between orphans and non-orphans is not large, or the prevalence of orphans is small, then the degree of the bias has little effect on the estimates of child mortality or their implications. In countries with high HIV prevalence and mortality, these assumptions may not hold. For this reason, IGME uses an adjustment for the bias in countries with high HIV prevalence.

IGME Bias Adjustment

Walker et al. (2012) describe the process of adjusting for bias due to AIDS in estimating under-5 mortality used by IGME.[17] They point out that the most accurate country-specific adjustments would require information about the distribution of births by duration of HIV infection, country-year specific rates of transmission of HIV from mother to child, and survival times of children and mothers after birth, among other things. Because these detailed data are largely unavailable, the IGME adjustments make several simplifying assumptions. The system uses an Excel workbook with the newest version of the UNAIDS Spectrum model results, which use estimates of HIV incidence and prevalence, as well as assumptions about mortality on and off of treatment, to produce births in three categories- HIV-negative births to HIV-negative mothers, HIV-negative births to HIV-positive mothers, and HIV-positive births to HIV-positive mothers. Then, the children and their mothers in this model are exposed to

mortality rates derived from cohort studies or model life tables. In the end, the model produces a proportion of births and child deaths that would have gone unreported because of a maternal death, which can be summed for any desired survey recall period to derive a correction factor. Figure 1 shows these categories and some of the assumptions described by Walker and colleagues to essentially generate the "true" number of births and deaths that wouldn't be captured by surveys because of mothers who have died due to HIV.

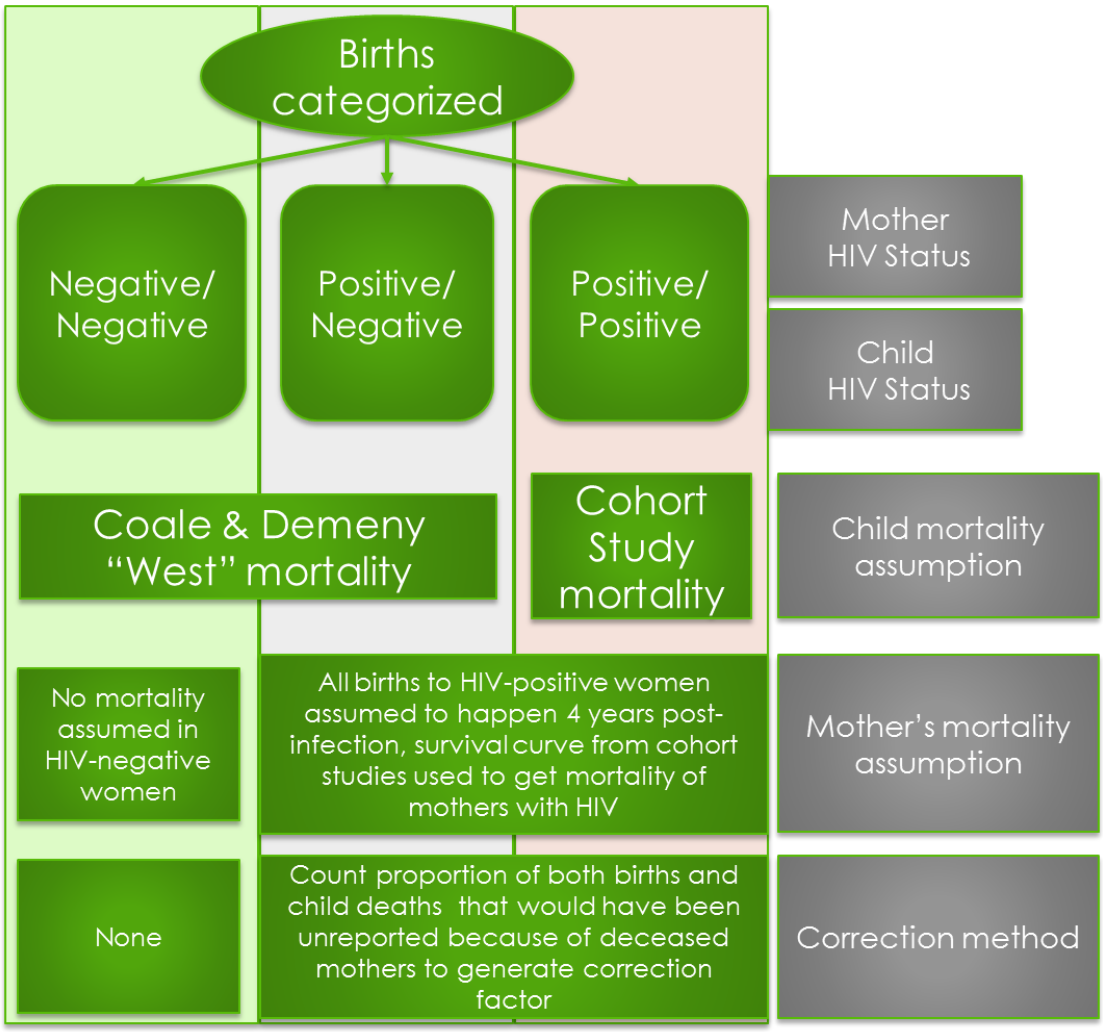


Figure 1: Outline of IGME Approach for HIV/AIDS Correction to CBH Methods

Alternative Data Sources

These correction methods used by IGME rely on generalizing cohort studies to other populations and making assumptions about HIV transmission. My analysis seeks to use nationally representative survey data as an alternative method to generating a correction factor. Specifically, orphan prevalence can be measured using the household modules from DHS surveys, and the ratio of orphan to non-orphan mortality rates can be estimated using sibling survival questions from DHS surveys.

Sibling survival questionnaires are commonly used in adult mortality estimation. Well-established and validated bias correction methods have been developed to account for selection bias in higher-mortality

sibships, sibships with zero surviving members, and recall bias.[25–28] However, literature on estimating child mortality using sibling survival is relatively limited. Masquelier (2013) and Merdad (2013) both note inconsistencies between estimates of under-5 mortality derived from sibling survival histories and complete birth histories, notably differences in the composition of birth orders between the sibling data and birth history data as well as recall bias differences.[29,30] I use several strategies to minimize the impact of these noted biases. This analysis estimates the degree of bias in complete birth history estimates using all available DHS surveys in Sub-Saharan Africa with household, sibling survival, and complete birth history modules.

Methods

Data

This analysis used data from the Demographic and Health Survey (DHS) program. The DHS surveys contain household questionnaires that inquire about the survival status of the parents of any child under fifteen years of age which can be used to estimate orphan prevalence. In addition, the DHS instrument contains modules with complete birth history questionnaires and sibling survival questionnaires, particularly for women. In total, 48 surveys in Sub-Saharan Africa between 1999 and 2014 contained all of the appropriate information to make the full set of estimates. In particular, to generate a full set of estimates, the specific DHS survey needs to have a sibling survival module, a household module that captures the survival status of each child's parents, and a birth history module. In some cases, the sibling survival module could not be used because the questions about parental survival in the household survey were only answered for children under the age of 15, so it was not possible to stratify responses to the sibling survival by orphan status. Several DHS surveys in Kenya and the South Africa DHS from 1998 were unusable because of the responses to that question. Tables in the supplement contain a list of surveys with descriptive information for the household and sibling survival portions used to estimate orphan prevalence and orphan mortality.

Orphanhood Prevalence

Many previous studies have used several different methods to estimate the prevalence of orphans. Hunter and Williamson (2000)[14] detailed several of the methods, including model-driven estimates from groups like UNAIDS or the United States Census Bureau and survey-driven estimates using sources like DHS and censuses. Each has its strengths and weaknesses. Modeling strategies are commonly used to estimate orphan prevalence due to AIDS, and theoretically, they would be able to estimate prevalence of unsheltered orphans or orphans in institutions who would not be captured in household surveys. However, while these projection models can be calibrated to estimate orphan prevalence using demographic and epidemiological inputs, they are not directly fitting orphan prevalence data.[16] In general, projection models that estimate all-cause and AIDS-specific deaths have been found to overestimate orphan prevalence, though continued methods development on models like the UNAIDS Spectrum HIV model are improving estimates with more detailed input data and assumptions.[31,32] New analyses continue to assess the quality of different sources of estimates of orphan prevalence and living conditions.[33]

In addition to the several strategies of estimating orphan prevalence, there are several quantities of interest in the literature. Various sources define orphans in different age ranges, with some estimates for children under 15 and others reporting for children under 18. Additionally, orphans are often described as maternal orphans, paternal orphans, and double orphans. This analysis is primarily focused

on maternal orphans, as these are the orphans not captured in birth history surveys. Here, maternal orphans are defined as children under 15 whose mothers are deceased.

To create a set of survey-specific estimates, orphan prevalence estimates were derived from household surveys to remain consistent with the survey-specific sibling survival and complete birth history estimates of child mortality, rather than using projection model estimates of orphan prevalence. The complete set of surveys used to estimate these prevalences were larger than the set used for the full analysis, as the only required component was a household module. However, there prevalences were only estimated in countries for which any sibling histories were available. A total of 1,989,103 children under 15 years of age, 708,715 of which were under age 5, were included in these 84 surveys, ranging from 5,914 children under 15 in São Tomé and Príncipe in 2008 to 80,356 in Nigeria in 2013. Microdata from each survey were extracted, and survey weights were used to tabulate the proportion of children under 15 and by single-year age groups within this range, with deceased mothers. Results were compared to estimates from the same surveys in other studies[31] and in DHS reports for validation. This analysis was conducted using the *survey* package[34] (version 3.30) in R (version 3.2.2). To propagate uncertainty in these estimates through to the rest of the analysis, 1000 draws of survey-age specific orphan prevalence were generated by sampling from a binomial distribution with the number of trials equal to the sample size in the age group in the survey.

Sibling Survival Method

Previous Methods

Sibling survival methods were initially developed as sisterhood methods for estimating maternal mortality in particular.[35] More recently, sibling survival questionnaires have become more commonly included in household surveys to create estimates of adult mortality. Methodology has evolved to correct for biases in the adult mortality estimates that arise from the sibling survival method, including the selection bias of selecting sibships with lower mortality risk in surveys because more siblings will be alive and the null probability of selecting a sibship in which all siblings have died.[25,26] These methods have continued to be corrected to improve their reliability, using sibling-specific weights when conducting analysis at the level of the sibling and creating a recall bias adjustment using previous surveys in the same country as a comparison.[27,28]

While sibling survival methods have largely been used for adult mortality estimation, they contain data that could be used for child mortality estimation. Few studies have sought to use sibling survival data to estimate child mortality. Previous analyses using sibling survival data to estimate child mortality have found challenges obtaining estimates consistent with those from birth histories. From sibling histories and birth histories taken in the same household, Merdad (2013) found that the sibling histories tend to report smaller family size (10% of the time) and fewer deaths (8% of the time) but are largely similar.[29] Overall differences in estimates of child mortality between the two methods exist, with sibling histories generally estimating higher mortality particularly in the years closer to the survey. One reason for this difference is the composition of the sample; the average birth order of children in the sibling analysis is higher. Indeed, when restricting analysis to children of certain birth orders, the sibling estimates of child mortality are lower than the birth history estimates, potentially because of omission from recall bias. Masqeulier (2013)[30] also found omission bias to be a large data quality issue in sibling histories, particularly in West Africa. However, the range of the bias is large. In the 2005 DHS in Rwanda, 2% of daughters report fewer siblings than their mother reported children, and 2% of mothers reported fewer

children than their daughters reported siblings. However, in the 2008 DHS in Sierra Leone, 20% of daughters reported fewer siblings than their mothers reported children, and 10% of mothers reported fewer children than their daughters reported siblings. In the 1999 Nigeria DHS, 22% of daughters reported fewer siblings than their mothers reported children, but this survey was dropped from this analysis due to data quality issues in line with what previous studies have done.[26]

Current Method

Despite these documented data quality concerns, this analysis utilized the sibling survival history method to estimate child mortality. However, I sought to minimize potential bias in several ways. First, I followed the work of Levin-Rector et al. (2013) in adjusting for selection bias from survivorship using established methods.[28] Second, while I used the whole set of time periods in the regression to estimate mortality using the sibling survival data, I only examined results from the 5 years prior to the survey to minimize the effect of recall bias. Third, as this analysis sought to estimate the ratio of orphan to non-orphan mortality for the purpose of quantifying the degree of selection bias in birth history methods, the only bias that affects these estimates would be bias that differentially affects orphans and non-orphans. Last, I have identified a source of this bias unique to this analysis, namely that person-years of exposure cannot be correctly classified as orphan and non-orphan without some assumptions about when a mother's death occurred because the exact year is unknown. I was unable to add a correction for zero-survivor sibships, as it is less clear how to categorize these sibships as orphans or non-orphans.

In the DHS, the sibling survival questionnaire can be paired with the household questionnaire to obtain sibling survival recall from 15- to 17-year olds, stratified by orphan status. Unfortunately, the household questionnaire does not ask the date of a parent's death, simply their survival status at the time of the survey, and the question is only answered among those younger than 18. Thus, person-years of exposure and deaths in the sibling survival analysis can be categorized as orphan or non-orphan, making some assumptions about the mother's year of death. The most naïve assumption is to categorize the whole sibship of an orphan respondent in the orphan group. A better assumption is to categorize all person-years of exposure before the birth of the youngest child as non-orphan, as the mother could not have died before the birth of her last child. I call this the baseline assumption. I have conducted sensitivity analyses to assess the degree to which the assumption about the date of the mother's death affects estimates of orphan and non-orphan mortality.

Figure 2 shows the effect of the baseline assumption that the mother died right after the last child was born. In each plot, three children in this example sibship are shown, born in 1986, 1989, and 1990. The second sibling dies in 1993, but the first and third children live to the time of the survey, at which point they are 16 and 12 years old, respectively. The top plot shows the baseline assumption, classifying the person-years of each child into years in which they are assumed to be orphans and non-orphans. The second plot shows a scenario in which the mother dies in 1992, two years after the birth of her last child. The baseline assumption classifies the years between 1990 and 1992 as orphan years; however, in this scenario these years should be classified as non-orphan years. This misclassification leads to an underestimate of orphan mortality. In the third plot, the mother dies in 1994, after her second child has died. The baseline assumption classifies this death as an orphan death, when it is really a non-orphan death, occurring before the death of the mother. This misclassification leads to an overestimate of orphan mortality.

To investigate the effects of the assumption about when the mother died, I explored multiple scenarios, with the assumption about the mother's death varied to 20%, 40%, 60%, and 80% of the time between the birth of the last child and the time of the survey. Results of this sensitivity analysis are described in the results and discussion, and supplemental figures are additionally shown.

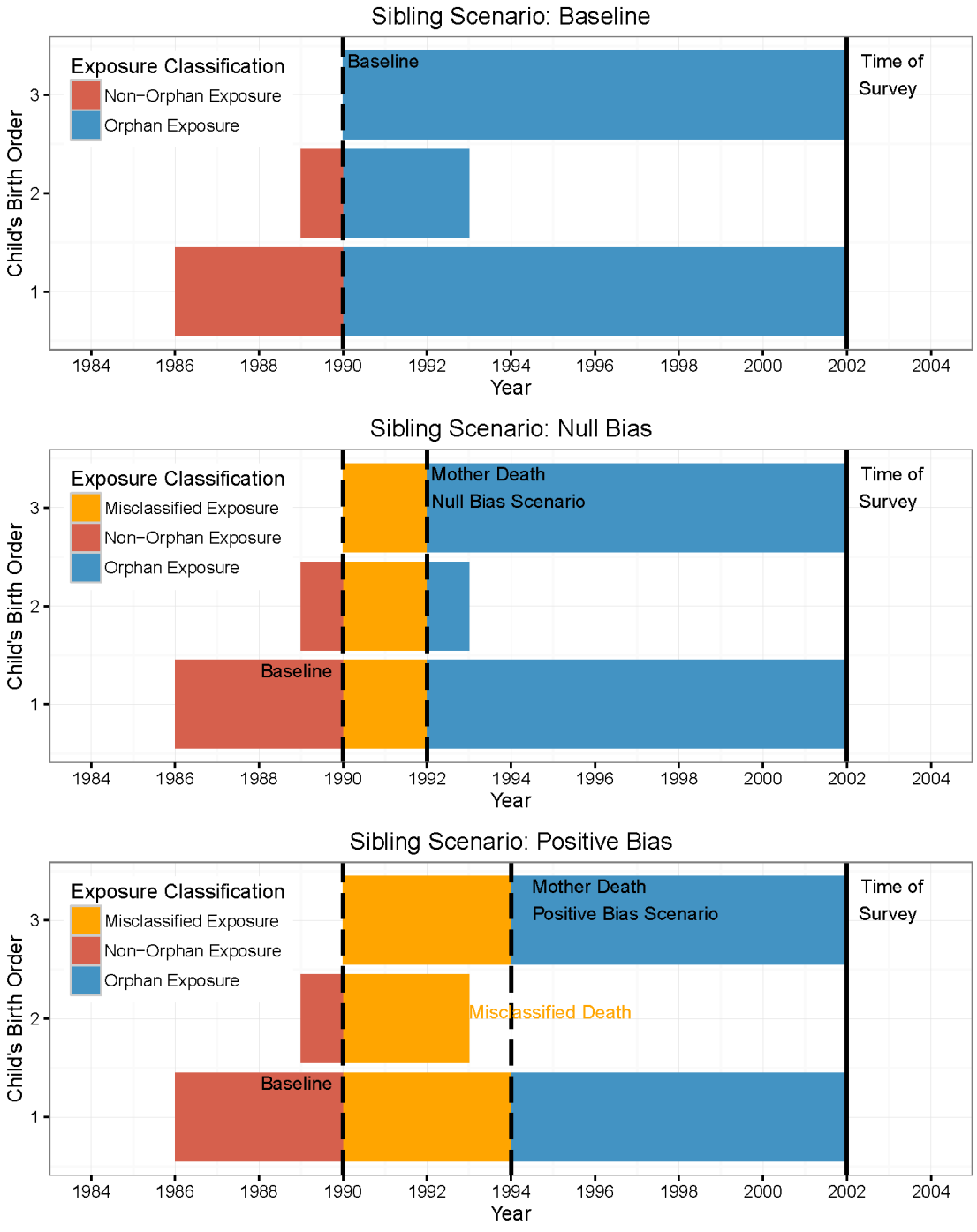


Figure 2: Sibling Maternal Death Assumption Scenarios

I separately estimated under-5 mortality in orphans and non-orphans by classifying deaths and exposure in the manner described, and estimating a logistic regression predicting deaths with indicators on age, with the combined sample weights and sibling weights, as well as accounting for the clustering in the survey design. To propagate uncertainty in my estimates, I created 1000 draws of my estimated probabilities of death for orphans and non-orphans for each age and time period of recall from each survey by sampling from the variance-covariance matrix of regression coefficients.[36] Then, in my analysis, I use the ratio of orphan to non-orphan child mortality from this process. This analysis was done using Stata version 13.1 and R versions 3.1.2 and 3.2.2.

Complete Birth History Method

The estimation of under-5 mortality using complete birth history data for all available DHS surveys has previously been done for the GBD 2015, and detailed descriptions of the methodology are available in Wang et al. (2016).[5] I used pooled 5-year estimates of survival probabilities in early neonatal (0-6 days), late neonatal (7-27 days), post-neonatal (1 month to 1 year), and single year age groups from ages 1 to 4 years derived from these surveys. To propagate uncertainty in these survey estimates to later stages of estimation, I created 1000 draws of survival by sampling from a binomial distribution, with the number of trials equal to the number of exposure months in the person-age.

Quantifying Bias in Birth History Estimates of Child Mortality

The underlying assumption in this analysis is that there is some factor by which estimates of child mortality from birth histories are biased due to the absence of deceased mothers in the sample. Sibling survival histories offer the opportunity to estimate the ratio of orphan to non-orphan mortality:

$$S_{i,t} = {}^{O}_{5}q_{0_{i,t}} / {}^{NO}_{5}q_{0_{i,t}}$$

where S is the ratio of orphan (*O*) to non-orphan (*NO*) ${}_5q_0$ from survey *i* as estimated for the 5-year period *t*. Each of these estimates of ${}_5q_0$ were derived from the sibling survival method described above. As others have shown, sibling survival estimates of under-5 mortality can be inconsistent with complete birth history estimates. However, some of the reasons for these issues, like sibling omission and compositional differences in birth order, may be less variable between orphans and non-orphans than between sibling and birth history estimates. Some potential problematic biases could still exist: differential omission bias between orphans and non-orphans and different ratios in orphan to non-orphan under-5 mortality among birth orders not as well represented in the survey, though sibling weights[25,26,28] should address some of this concern.

This scalar of orphan to non-orphan mortality, $S_{i,t}$, was then applied to complete birth history estimates of ${}_5q_0$ from the same survey *i* for the same group of years *t* to obtain an estimate of orphan ${}_5q_0$:

$${}^{"Orphan"\ CBH}_{5}q_{0_{i,t}} = S_{i,t} * {}^{CBH}_{5}q_{0_{i,t}}$$

Each of the probabilities of death in the under-5 age groups was then rescaled to aggregate to ${}_5q_0$ using conditional probabilities. These probabilities of death were then aggregated to single-year age groups: ${}_1q_0$, ${}_1q_1$, ${}_1q_2$, ${}_1q_3$, and ${}_1q_4$ and converted to m_x space using standard demographic assumptions.[9] Then, to estimate the degree to which selection bias of alive mothers affects complete birth history estimates, the orphan and non-orphan mortality rates were weighted using the prevalence of orphans and non-orphans:

$$^{Adjusted}_{1}m_x = P_{Orphans}\left(^{"Orphan"\ CBH}_{1}m_x\right) + P_{Non-Orphans}\left(^{CBH}_{1}m_x\right)$$

where P is the prevalence of orphans and non-orphans. This calculation was done by single-year age groups from 0 to 4. Each of these adjusted m_x values, weighted to include both orphan and non-orphan mortality, were then converted back to probability space using the same demographic assumptions and aggregated to ${}_5q_0$, resulting in an under-5 mortality estimate adjusted for selection bias due to orphanhood. The ratio of this corrected ${}_5q_0$ to the uncorrected complete birth history estimate of ${}_5q_0$ is referred to here as the bias adjustment factor. Uncertainty was maintained through all of these calculations by randomly pairing the aforementioned 1000 draws of each quantity in each calculation with the other quantities, and the means and 2.5th and 97.5th percentiles were used as summary metrics.

Results

Maternal Orphan Prevalence

The highest observed under-15 prevalence of maternal orphans, 9.7% (9.3%-10.1%), was observed in Lesotho in 2009, while the lowest, 1.1% (0.8%-1.4%), was observed in São Tomé and Príncipe in 2008. In total, 17 country-years out of 82 surveyed from 1992 to 2014 produce estimated prevalence of maternal orphans under 15 years of age to be above 5%. Those high estimates of prevalence occur in Lesotho in 2004, 2009, and 2014; Malawi in 2000 and 2004; Mozambique in 1997; Namibia in 2006; Rwanda in 2000 and 2005; Swaziland in 2006; Uganda in 1995, 2000, and 2006; Zambia in 2001 and 2007; and Zimbabwe in 2005 and 2010. Figure 3 shows maternal orphan prevalence in Rwanda and Zimbabwe by 5-year age group and for children under the age of 15. The lasting effects of the Rwanda genocide on orphan prevalence in higher age groups is evident in 2000, but declines drastically with newer cohorts of children being captured by the 2010 DHS. Meanwhile, in Zimbabwe, the HIV epidemic took a large toll, with the prevalence of maternal orphans in ages 10 to 14 rising from under 5% in 1994 to over 15% in 2010.

Prevalence of orphans under the age of 5 is much lower. By the time children are 15, mothers have been exposed to 10 additional years of mortality risk compared to mothers when children are 5 years old, so the difference is unsurprising. The highest observed under-5 prevalence of maternal orphans, 2.4% (2.0%-2.8%), was observed in Zimbabwe in 2005, while the lowest, 0.1% (<0.1%-0.3%), was observed in São Tomé and Príncipe in 2008. In total, 12 country-years out of 82 surveyed estimated the prevalence of maternal orphans under 5 years of age to be above 1.5%, including Eritrea in 1995; Lesotho in 2004, 2009, and 2014; Rwanda in 2000; Sierra Leone in 2008 and 2013; Swaziland in 2006; Uganda in 1995; and Zimbabwe in 1999, 2005, and 2010. See the supplemental tables for prevalence of maternal orphans under the age of 15 and under the age of 5 for the set of surveys analyzed.

Orphan and Non-Orphan Mortality

The estimation of orphan and non-orphan mortality using sibling survival methods largely shows higher mortality among orphans. This ratio was significantly larger than 1 in 20 out of 48 surveys. Five surveys- Benin in 2006, Côte d'Ivoire in 2011, Congo in 2011 Gabon in 2012, and Swaziland in 2006 had estimates of the orphan to non-orphan under-5 mortality ratio lower than 1, though they were all insignificant. From the Zimbabwe 2010 survey, the under-5 mortality in orphans reached 9.69 (4.0 to 18.5) times higher than the mortality among non-orphans, and 9 surveys had mean estimates of orphan mortality 4 times higher than non-orphan mortality. These surveys occurred in Congo in 2005; Liberia in 2006;

Lesotho in 2004; Mali in 2012; Namibia in 2006 and 2013; Togo in 2013; and Zimbabwe in 2005 and 2010.

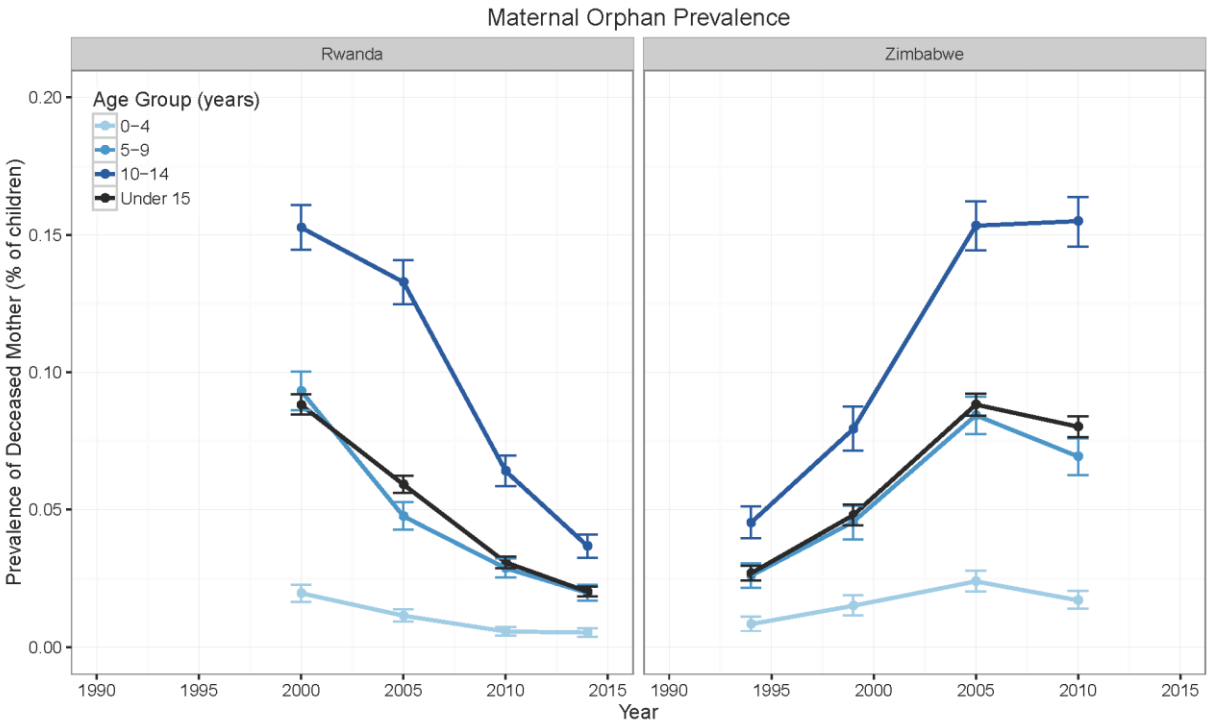


Figure 3: Maternal Orphan Prevalence in Rwanda and Zimbabwe, by Age Group

One concern about the validity of the sibling survival results in this analysis is the sample size. The number of respondents ranged from 262 to 4,735 15 to 17 year-olds interviewed in Guinea in 1999 and Nigeria in 2013, respectively. These same two surveys included a total of 861 and 13,676 person-years as well as 23 and 315 deaths among all of the respondents' siblings under the age of 5 in the 5 years prior to the survey, with 36 and 463 person-years as well as 2 and 13 deaths among orphans in the 5 years before these surveys. Across the 48 surveys, an average of 2234 person-years were captured in the 5 years prior to the survey in the under-5 age group, 158 of which were classified as orphan years under the baseline assumption. Meanwhile, a mean of 46 deaths were captured, with 7 classified as orphan deaths under the baseline assumption. The supplemental tables show descriptive statistics about the samples in each of the components of the surveys used in this analysis. Larger sample sizes would improve the quality of these estimates. For example, the Swaziland 2006 survey estimates an orphan to non-orphan mortality ratio under 1, but that is based on only 1 orphan death captured in the sibling recall during the 5 years prior to the survey (though data before that time period also informs the estimate to some degree). On the other hand, some surveys have larger samples and more stable results. Panel 1 of Figure 4 shows the ratio of orphan to non-orphan mortality for all analyzed surveys.

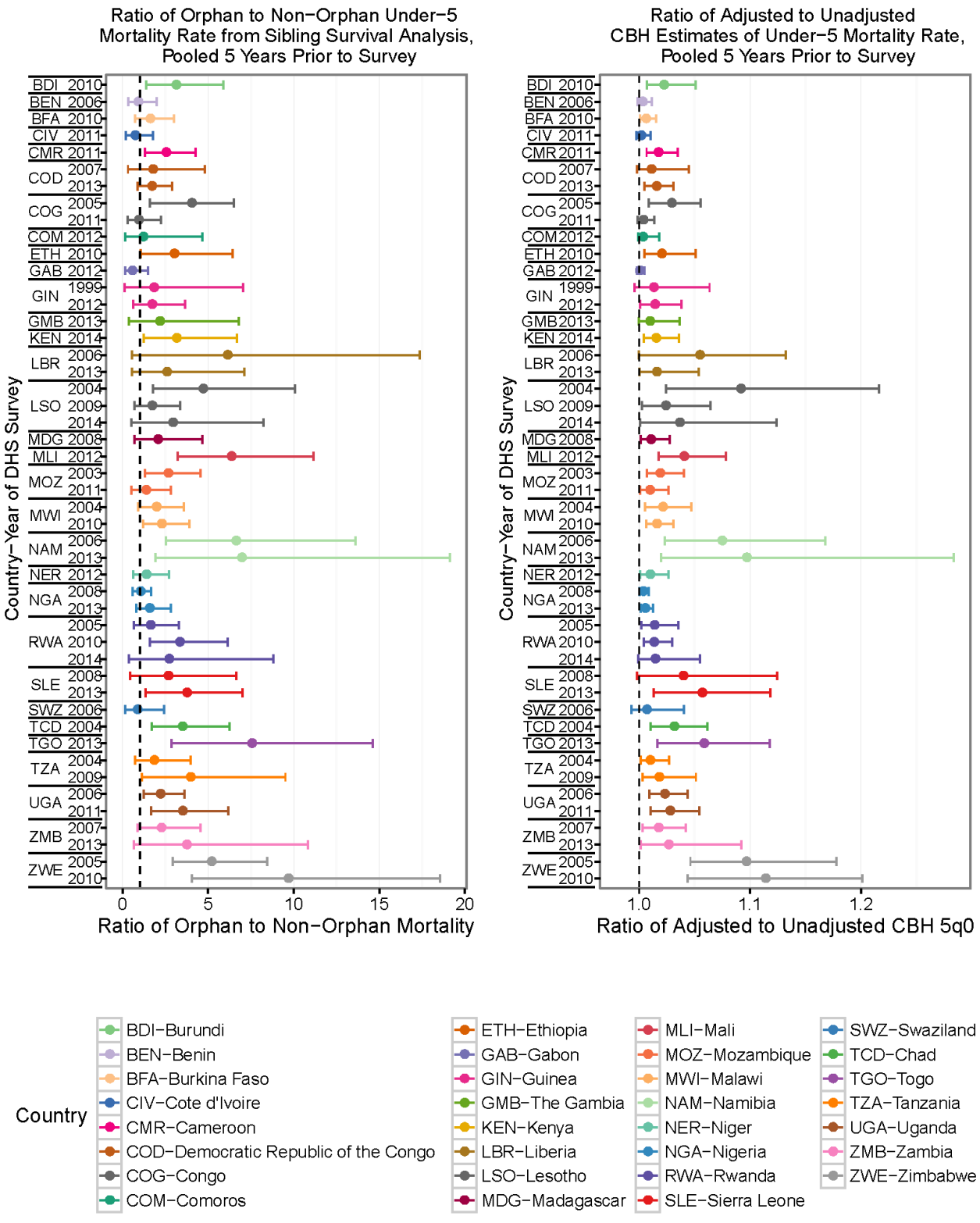


Figure 4: Ratio of Orphan and Non-Orphan Under-5 Mortality Rate from Sibling Survival Analysis and Ratio of Adjusted to Unadjusted CBH Estimates of Under-5 Mortality Rate

One of the other concerns regarding the quality of the sibling survival estimates of under-5 mortality relates to the assumption of when a mother has died. For the purposes of the full analysis, I used the assumption that the mother died just after the birth of the youngest child. However, I tested the effect of the assumptions that she died after 20%, 40%, 60%, or 80% of the time between the birth of her last child and the time of the survey. The drastic reduction in the number of person-years and deaths in the orphan group if 80% of the time after the birth of the youngest child is reclassified as non-orphan creates challenges with estimation. Results become unstable, and in some cases, zero deaths are then observed, so the ability to estimate results for more surveys is lost with higher assumptions. However, the effect of increasing the assumption from 0% to 20% or 40% had mixed effects. The estimates of orphan to non-orphan mortality increased in some surveys and decreased in others, as both deaths and person-years of exposure became reclassified from orphan to non-orphan. Plots of the effects of the different assumptions are shown in the supplementary figures.

Quantifying Bias in Birth History Estimates of Child Mortality

The final adjusted under-5 mortality rate using both the prevalence of orphans and the ratio of orphan to non-orphan mortality to recover the all-child mortality rate was statistically significantly higher than the unadjusted CBH estimate in 36 out of 48 surveys. Figure 4 shows the ratio of the adjusted to unadjusted estimates of under-5 mortality on the panel on the right. The highest correction estimates were found in Zimbabwe, Namibia, and Lesotho. The Zimbabwe 2010 DHS adjusted child mortality for the 5 years prior to the survey were 11.4% higher (4.4% to 20.0%) than the unadjusted estimates. Meanwhile, 20 surveys had adjustment factors more than 2%, and only 8 surveys had an adjustment factor more than 5% suggesting that in most cases, the effect that selection bias has on CBH estimates is relatively small. Overall, this method makes small but significant adjustments to most surveys in Sub-Saharan Africa.

Both the maternal orphan prevalence and the ratio of orphan to non-orphan under-5 mortality effect the size of the difference between the final adjusted under-5 mortality rates and those derived from the complete birth histories. Figure 5 shows how both of these factors interact to affect the ratio of the adjusted estimates to the unadjusted estimates of under-5 mortality. Some surveys have quite similar bias adjustment factors for largely different reasons. For example, the prevalence of maternal orphans in Zimbabwe in 2005 was quite high compared to that of Namibia in 2013; however, the two surveys have similar adjustment factors because the ratio of orphan to non-orphan mortality was higher from the Namibia survey. This figure illustrates the range of each of the three parameters across all of the surveys. It also shows some instability in the estimates from the sibling survival method. For example, the ratio of orphan to non-orphan mortality shifts between almost 5, less than 2, and about 3 between the 2004, 2009, and 2014 surveys in Lesotho. While some shifts are certainly plausible given changing social factors and differences in stages of the HIV epidemic, such shifts seem quite abrupt. Of course, when visualized with uncertainty, as in Figure 4, it becomes clear that this decrease and then increase over time is not statistically significant and there is a large amount of uncertainty in most of these adjustments. However, there is also a degree of consistency. Zimbabwe, Sierra Leone, and Namibia clearly have among the highest adjustment factors for several of their surveys, with relatively more consistent estimates of orphan to non-orphan mortality in Namibia and Sierra Leone over time.

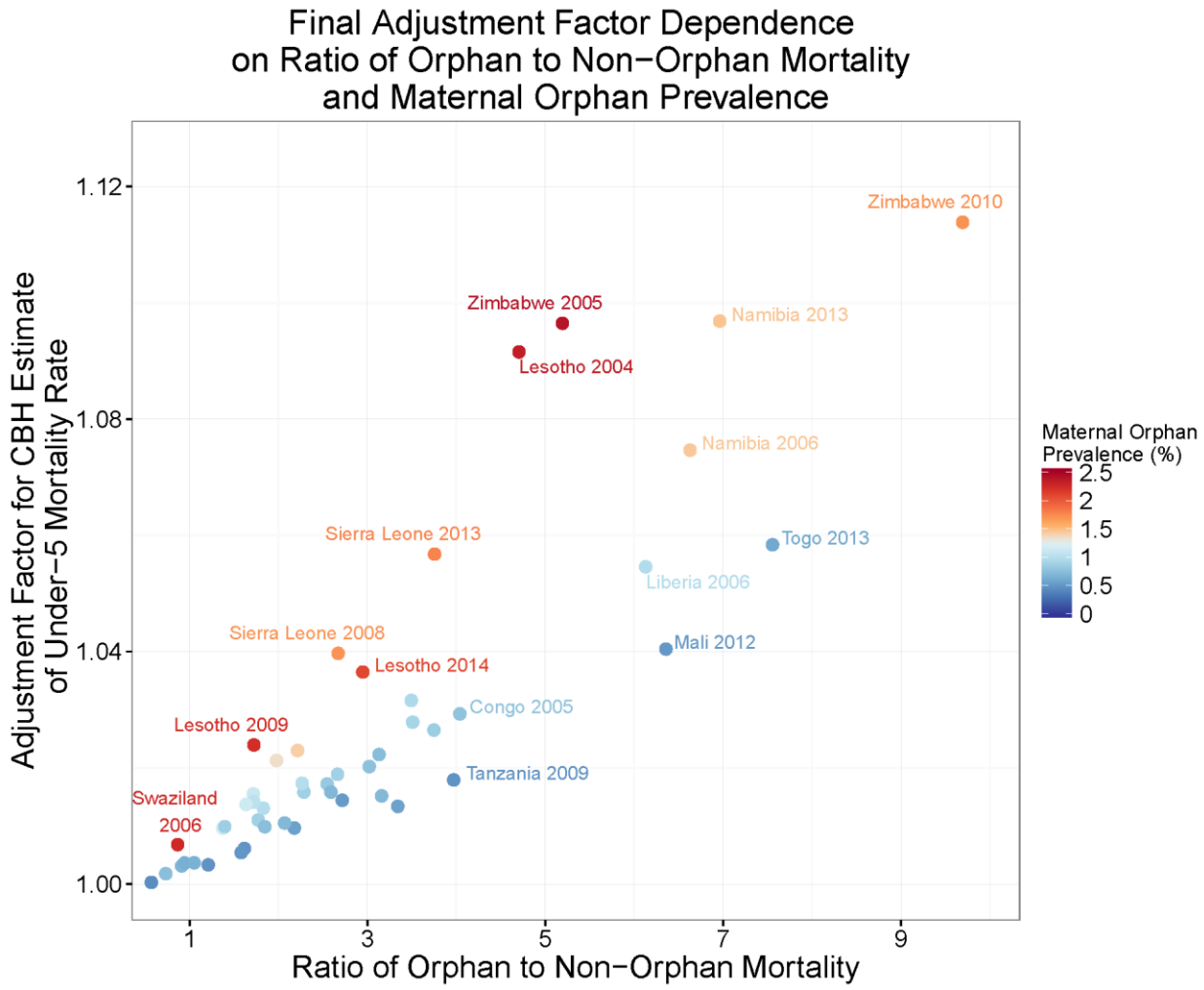


Figure 5: Effect of Maternal Orphan Prevalence and the Ratio of Orphan to Non-Orphan Mortality on the Adjustment Factor for CBH Estimates

Discussion

Quality of Results

This method of analysis seeks to offer an alternative process of correcting for selection bias in complete birth history estimates of under-5 mortality using survey data from the same surveys as the birth history modules. It allows the use of mortality data specific to the country of interest and estimates bias due to orphan mortality because of all causes rather than an AIDS-specific correction. The findings did show statistically significant, though often small bias in many complete birth history estimates of child mortality. In some ways, the method produces consistent estimates; the countries with the highest bias adjustment factors have multiple surveys indicating that they have a large amount of bias. However, some of the patterns within countries are somewhat difficult to explain. The orphan to non-orphan mortality ratio in Zimbabwe doubled between the 2005 and 2010 surveys. Although Lesotho stood out as a country that had consistently high prevalence of orphans and a high adjustment factor, the size of the adjustment varied substantially.

Several factors contribute to the instability. First, the sample sizes in the portion of the sibling survival modules that could be used for this analysis are relatively small compared to the household sample sizes that could be used to inform the estimates of orphan prevalence. This is reflected in the large uncertainty intervals around the estimates of orphan to non-orphan under-5 mortality rates. Additionally, assumptions about when the mother of an orphaned sibship dies between the time of the last born child and the time of the survey may play a role in the instability. Across all the surveys, reclassifying the first 20% of time between the birth of the last child and the time of the survey among the orphan group from "orphan" to "non-orphan" did not have much of a net effect; however, it did have potentially large consequences on individual surveys. I sought here to create an adjustment factor that relied solely on modules all from a given survey, but the assumptions made about when mothers die in the orphan group might improve if other data can be used to estimate the mean number of years lived in that interval. For example, making some informed assumptions about the mother's age given the number of children she had, the mean number of years lived between the year the last child was born and the time of the survey could be estimated from sources like the GBD study by finding the mean age at death among women who died in the corresponding cohort between those years. However, using this assumption could be infeasible if it would reclassify too many of the remaining orphan person-years of exposure and deaths, shrinking the sample size significantly.

Despite the opportunity for future methods development, this method found significant bias in the estimates of under-5 mortality from complete birth histories. To further validate the estimates, they can be compared to the existing IGME bias adjustment factors, primarily driven by HIV estimation.

Comparison to IGME Adjustments

While we might expect somewhat different adjustments than the IGME method, they should be similar because HIV/AIDS is such a large contributing factor to orphan prevalence and mortality, particularly in Sub-Saharan Africa. Figure 6 shows a scatter plot comparing IGME adjustments from their 2015 version of their estimations[4] with those from this method. Included are surveys for which both methods generated adjustments as well as surveys for which only one of the methods generated adjustments (shown as a ratio of 1 if no adjustment was generated). To approximate the IGME data adjustments over the 5 years prior to the survey, I downloaded their adjusted and unadjusted data from their website (www.childmortality.org) and calculated the adjustment factor for the estimates in each set of pooled years in which their estimates are generated (different than the 5-year pooling that this method is using). Then, I interpolated between years to create estimates for every 6 months. To create a pooled estimate in the 5 years prior to the survey, I simply took the average of the adjustments in that 5-year span. While this approximation is not an exact calculation of what the 5-year pooled adjustment would be, it is sufficient for comparison between the two methods.

There is consistency in the adjustment factors between the two methods in a number of surveys, including Zimbabwe 2010, Lesotho 2004, Namibia in 2006 and 2013, and several surveys with lower adjustments like Zambia 2013. However, some other surveys clearly show discrepancies. For example, the Swaziland 2006 survey gets a very large correction from the IGME method, but essentially no correction from the current method. This survey was pointed out earlier as an example of an unstable estimate due to the small number of deaths captured. In addition to instability, our results differ in the choice of which surveys to correct. While the IGME method can correct for surveys that this method cannot for lack of sibling survival modules, they choose to apply their method to a certain subset of locations with high HIV prevalence. My results show a potential limitation in this decision; some of the

country-years not included in the IGME correction locations show bias in my results. Togo in 2013, Sierra Leone in 2013, and Liberia in 2006 all have over a 5% bias correction according to my method, while the IGME estimates do not make any correction. On the other hand, this method was unable to find adjustments for Zimbabwe 1999, South Africa 2004, Kenya 2008, and Kenya 1998 because the proper data were not available, and they have quite large adjustment factors based on the IGME method. This limitation also points to the utility of cross-survey regressions or pooling to estimate adjustments for surveys without all of the required data for this correction.

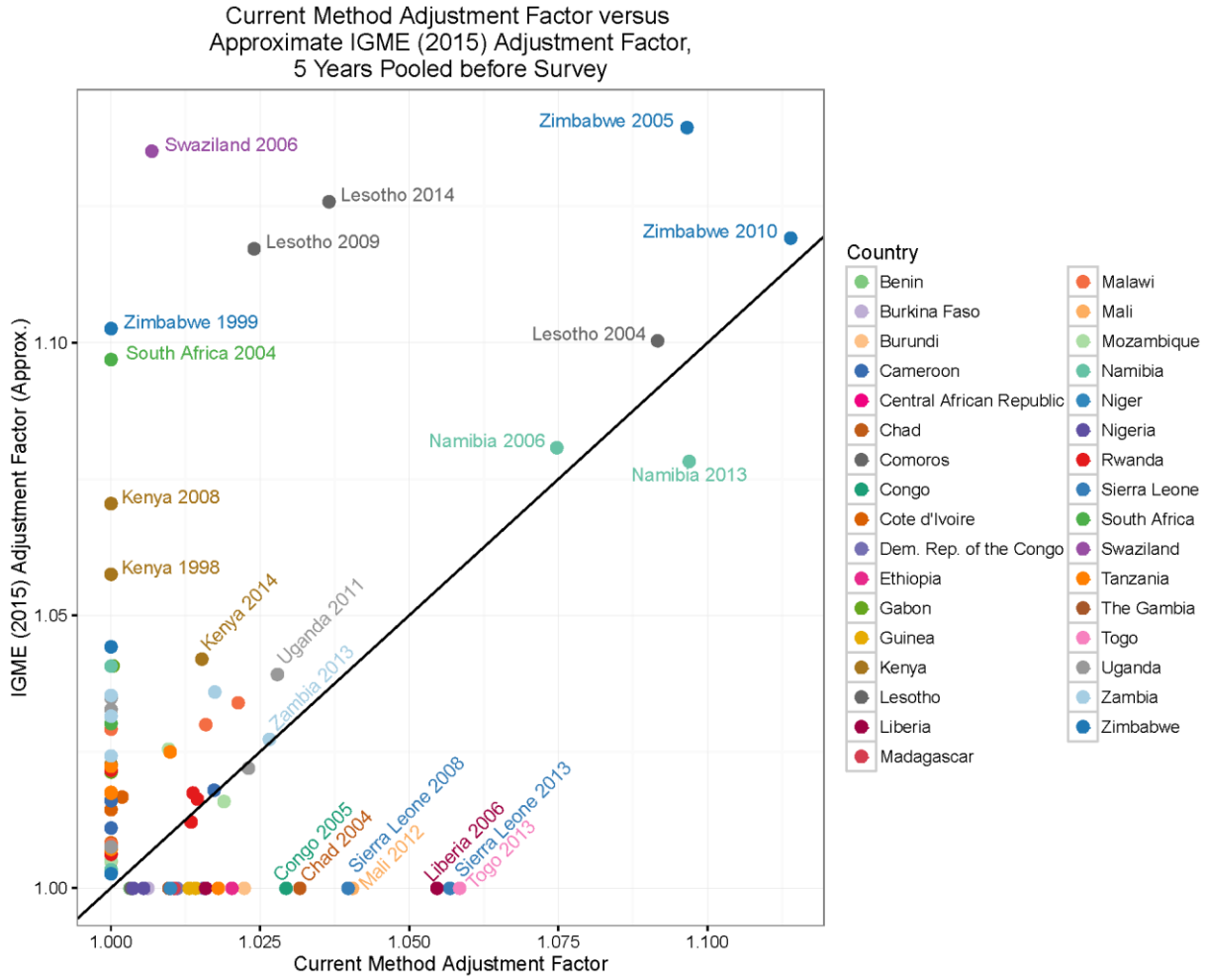


Figure 6: Comparison of Current Method to IGME Spectrum-based Adjustment Method

Limitations

This novel correction method has several strengths. Using nationally representative survey data does not force assumptions to be made about the representativeness of epidemiological estimates from cohort studies about transmission of HIV or survival time. It also allows for the estimation of the effect of the bias from all causes rather than focusing on HIV. However, there are several limitations to the method, particularly with respect to the sibling survival estimates of under-5 mortality.

First, there are limitations in sample size. As shown in the supplemental tables, the sample of 15-17 year olds who answer both questions about the survival of their parents and the sibling survival module is

relatively small. In some surveys, it is adequate to produce reasonably reliable estimates; however, it is clear that in some surveys, estimates are unstable because of the small sample size. While this analysis focused on a survey-specific adjustment derived from one survey, future analyses could help address the stability of the estimates by analyzing all available surveys simultaneously, drawing predictive strength from covariates and data from similar geographic locations or time periods.

Second, several biases are problematic in sibling survival estimates of under-5 mortality. Recall bias is a concern, as children may not remember the death of their older siblings if they were young when the event occurred, and they may simply not be as reliable at remembering details about siblings as mothers are at remembering details about their children. This study only reported estimates from the 5-year period before the survey in order to mitigate this bias. Future analyses may be able to utilize estimates from earlier periods by deriving recall bias corrections from comparing to previous surveys in the same country when available, as is done to estimate adult mortality.[28] Additional information may also be gained by comparing responses in mothers and daughters from the same survey as others have done[30] in order to try to derive corrections for sibling recall. This method, however, would be limited by small numbers. Compositional bias is also a concern. Merdad documented differences in the breakdown of birth orders among children considered in analysis between CBH and sibling survival samples.[29] I have sought to minimize this bias by simply calculating a ratio of orphan to non-orphan mortality with both estimate coming from sibling survival methods. This eliminates the bias in comparing CBH estimates to sibling survival estimates; however, if the sibling survival samples do not represent the true proportion of birth orders in the population, then further adjustment might be needed. Finally, the lack of information about when the mother of an orphaned sibship died may affect the results. Sensitivity analysis showed that this assumption can change estimates substantially on a survey-to-survey basis, although there is not a clear overall direction of this bias. Further methods to address this concern should be explored.

The timing of the estimates of all of the parameters is also a limitation. The estimates of maternal orphan prevalence are at the time of the survey, while the mortality estimates are based on recall, and therefore placed in time a couple of years before the survey. The HIV epidemic has a very large effect on orphan prevalence, so changes of two or three years may have a measurable impact on the prevalence of orphans. In the future, the estimation of a complete time series of orphan prevalence would allow matching mortality estimates from any point in time with a more appropriate prevalence of orphans. Such an analysis could simply interpolate orphan prevalence if enough survey data are available, or use covariates such as maternal mortality estimates from the GBD to predict orphan prevalence. This approach was not taken, here, as the goal was to derive an adjustment factor strictly from the survey of interest. Even with a time series of orphan prevalence estimates generated from surveys, there are children missed because they live in institutional settings or are unsheltered. This may cause an underestimate in the bias adjustment not only because they are not included in the prevalence of orphans, but also because they may be at higher risk of mortality.

Future Directions

This analysis shows a step towards empirical estimates of bias adjustment factors for complete birth history estimates of child mortality due to selection bias of mothers who are still alive. Several methodological improvements suggested in the limitations include further adjustment to sibling survival methodology, specifically for estimating child mortality. These include recall bias correction improvements, adjustments for compositional bias, and more informed assumptions about the timing of

the deaths of mothers of orphaned sibships. In addition to some of these methodological improvements, alternative ways of validating this analysis like simulations would improve confidence in the accuracy of the results.

Part of the advantage of this method is that it does not rely on projection models and can estimate an adjustment factor based on orphanhood from all causes. Applying this methodology to surveys outside of Sub-Saharan Africa, particularly in countries with and without high prevalence of orphans, could show whether this selection bias in complete birth histories may be a problem more widespread than simply in countries or time periods with high HIV prevalence. The expansion of the analysis to new geographies and beyond a survey-by-survey basis would offer more opportunity to use the results. Correction factors could be estimated in DHS surveys that do not contain sibling survival histories by creating comprehensive estimates of orphan prevalence and ratios of orphan to non-orphan mortality. These correction factors could then be applied across a range of surveys, potentially influencing the final estimates of child mortality in the Global Burden of Disease studies.

This analysis focused on under-5 mortality, but the findings also have implications on estimates of older child and adolescent mortality. Although the ratio of orphan to non-orphan might be lower in children aged 5 to 15 than in children under the age of 5 because they are less dependent on their mothers, the prevalence of orphans over 5 is much higher. An assessment of the bias in complete birth history estimates of mortality between ages 5 and 15 using this method would likely find meaningful results.

Conclusions

This study found statistically significant and non-negligible adjustment factors to correct complete birth histories for selection bias. While the estimates of bias are smaller than 5% in most cases, there are cases in which estimates of child mortality would be meaningfully changed by correcting for the bias. Although discrepancies exist between the results of this method and that of IGME, there is a general concordance in the estimates. Further validation and methodological development need to continue for this method to be used as an alternative for estimating correction factors.

References

1 Hauser PM. Demographic Indicators of Economic Development. *Econ Dev Cult Change* 1959; **7**: 98–116.

2 Sen A. Mortality as an Indicator of Economic Success and Failure. *Econ J* 1998; **108**: 1–25.

3 United Nations. United Nations Millennium Declaration. 2000; published online Sept 18. http://www.un.org/millennium/declaration/ares552e.pdf.

4 You D, Hug L, Ejdemyr S, *et al.* Global, regional, and national levels and trends in under-5 mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Inter-agency Group for Child Mortality Estimation. *The Lancet* 2015; **386**: 2275–86.

5 GBD 2015 Child Mortality Collaborators. Global, regional, national, and selected subnational levels of stillbirths, neonatal, infant, and under-5 mortality during 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet* In review.

6 UN General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development. UN General Assembly, 2015 https://sustainabledevelopment.un.org/post2015/transformingourworld (accessed Aug 11, 2016).

7 Hill K. Approaches to the Measurement of Childhood Mortality: A Comparative Review. *Popul Index* 1991; **57**: 368–82.

8 Silva R. Child Mortality Estimation: Consistency of Under-Five Mortality Rate Estimates Using Full Birth Histories and Summary Birth Histories. *PLOS Med* 2012; **9**: e1001296.

9 Preston S, Heuveline P, Guillot M. Demography: Measuring and Modeling Population Processes, 1 edition. Malden, MA: Wiley-Blackwell, 2000.

10 Ronsmans C, Chowdhury ME, Dasgupta SK, Ahmed A, Koblinsky M. Effect of parent's death on child survival in rural Bangladesh: a cohort study. *The Lancet* 2010; **375**: 2024–31.

11 Hallett TB, Gregson S, Kurwa F, *et al.* Measuring and correcting biased child mortality statistics in countries with generalized epidemics of HIV infection. *Bull World Health Organ* 2010; **88**: 761–8.

12 Nakiyingi JS, Bracher M, Whitworth JA, *et al.* Child survival in relation to mother's HIV infection and survival: evidence from a Ugandan cohort study. *AIDS Lond Engl* 2003; **17**: 1827–34.

13 Mahy M. Measuring child mortality in AIDS-affected countries. New York: UN Population Division, 2003 http://www.un.org/en/development/desa/population/events/pdf/expert/5/UNICEF_Paper15.pdf (accessed Aug 9, 2016).

14 Hunter S, Williamson J. Children on the Brink: Updated Estimates & Recommendations for Intervention. Washington, D.C.: United States Agency for International Development, 2000 http://pdf.usaid.gov/pdf_docs/pnacj284.pdf (accessed July 30, 2016).

15 UNAIDS, UNICEF, USAID. Children on the Brink 2004: A Joint Report of New Orphan Estimates and a Framework for Action. http://www.unicef.org/publications/cob_layout6-013.pdf (accessed July 31, 2016).

16 Stover J. Projecting the demographic consequences of adult HIV prevalence trends: the Spectrum Projection Package. *Sex Transm Infect* 2004; **80**: i14–i18.

17 Walker N, Hill K, Zhao F. Child Mortality Estimation: Methods Used to Adjust for Bias due to AIDS in Estimating Trends in Under-Five Mortality. *PLOS Med* 2012; **9**: e1001298.

18 Crampin AC, Floyd S, Glynn JR, *et al.* The long-term impact of HIV and orphanhood on the mortality and physical well-being of children in rural Malawi. *AIDS Lond Engl* 2003; **17**: 389–97.

19 Ward P, Zaba B. The Effect of HIV on the Estimation of Child Mortality Using the Children Surviving/Children Ever Born Technique. *South Afr J Demogr* 2008; **11**: 39–73.

20 Musariri H. The effect of HIV/AIDS on the own-child method of estimating child mortality: Lesotho and Zimbabwe. 2011. https://open.uct.ac.za/handle/11427/12751?show=full.

21 Nduru PM. Investigating biases in census questions on mortality using Agincourt health and demographic surveillance system data. 2013; published online Nov. http://open.uct.ac.za/handle/11427/5891?show=full.

22 Ayalew KA. Assessment of the IGME methods of estimating infant mortality rate and neonatal mortality rate from under-five mortlaity rate in countries affected by HIV. 2012. https://open.uct.ac.za/handle/11427/5896?show=full.

23 You D, Wardlaw T, Salama P, Jones G. Levels and trends in under-5 mortality, 1990-2008. *Lancet Lond Engl* 2010; **375**: 100–3.

24 Rajaratnam JK, Marcus JR, Flaxman AD, *et al.* Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: a systematic analysis of progress towards Millennium Development Goal 4. *The Lancet* 2010; **375**: 1988–2008.

25 Gakidou E, King G. Death by survey: Estimating adult mortality without selection bias from sibling survival data. *Demography*; **43**: 569–85.

26 Obermeyer Z, Rajaratnam JK, Park CH, *et al.* Measuring Adult Mortality Using Sibling Survival: A New Analytical Method and New Results for 44 Countries, 1974–2006. *PLOS Med* 2010; **7**: e1000260.

27 Masquelier B. Adult mortality from sibling survival data: a reappraisal of selection biases. *Demography* 2013; **50**: 207–28.

28 Levin-Rector A, Wang H, Rajaratnam JK, Levitz C, Schumacher A, Murray CJ. Improved analysis of sibling survvival data taking into account survivor bias, zero-surviving reporters and recall bias. In: Indirect estimation of mortality assessment and improvement of methods. Busan, South Korea, 2013. http://iussp.org/sites/default/files/event_call_for_papers/Sibling%20survival%20paper.pdf (accessed July 31, 2016).

29 Merdad L. Investigating New Methods for Estimating Maternal and Child Mortality in the Developing World. 2013.

30 Masquelier B. Inconsistencies in reports on siblings and children in survey data used to estimate mortality. In: Data quality in demographic surveys. Busan, South Korea, 2013.

31 Grassly N, Lewis J, Mahy M, Walker N, Timæus I. Comparison of household-survey estimates with projections of mortality and orphan numbers in sub-Saharan Africa in the era of HIV/AIDS. *Popul Stud* 2004; **58**: 207–17.

32 Stover J, Andreev K, Slaymaker E, *et al.* Updates to the Spectrum model to estimate key HIV indicators for adults and children. *AIDS Lond Engl* 2014; **28**: S427–S434.

33 Shoko M, Ibisomi L. Orphanhood Prevalence, Living Arrangements and Orphanhood Reporting in Lesotho, Malawi and Zimbabwe. *Child Indic Res* 2016; : 1–15.

34 Lumley T. Survey: Analysis of complex survey samples. *R Package Version 330*.

35 Graham W, Brass W, Snow RW. Estimating maternal mortality: the sisterhood method. *Stud Fam Plann* 1989; **20**: 125–35.

36 King G, Tomz M, Wittenberg J. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *Am J Polit Sci* 2000; **44**: 347–61.

Supplementary Figures

Supplemental Table: Maternal Orphan Prevalences Estimated from DHS Household Modules and Sample Sizes				
	Under 15 Years of Age		Under 5 Years of Age	
DHS Survey	Number of Children in Survey	Maternal Orphan Prevalence	Number of Children in Survey	Maternal Orphan Prevalence
Benin 1996	13746	2.5% (2.2%,2.7%)	5061	0.6% (0.4%,0.9%)
Benin 2006	44641	2.3% (2.2%,2.5%)	16405	0.7% (0.6%,0.8%)
Burkina Faso 2003	28666	2.7% (2.5%,2.8%)	10099	0.7% (0.5%,0.8%)
Burkina Faso 2010	39827	1.7% (1.6%,1.8%)	14466	0.5% (0.4%,0.6%)
Burundi 2010	19477	3.9% (3.6%,4.2%)	7683	0.7% (0.5%,1%)
Cameroon 2011	32321	3.2% (3%,3.3%)	12078	0.8% (0.7%,1%)
Central African Republic 1994	12853	3.9% (3.6%,4.2%)	4601	1% (0.8%,1.3%)
Chad 2004	14501	2.9% (2.6%,3.2%)	5347	1% (0.7%,1.2%)
Comoros 2012	9858	1.3% (1.1%,1.5%)	3349	0.5% (0.3%,0.7%)
Congo 2005	13230	3.2% (2.9%,3.5%)	4944	0.8% (0.6%,1%)
Congo 2011	23423	2% (1.8%,2.1%)	9545	0.6% (0.5%,0.8%)
Cote d'Ivoire 1994	17979	2% (1.8%,2.2%)	6400	0.5% (0.4%,0.7%)
Cote d'Ivoire 2011	22441	2.9% (2.7%,3.2%)	8155	0.7% (0.6%,1%)
Democratic Republic of the Congo 2007	22772	3.1% (2.9%,3.3%)	8803	0.9% (0.7%,1.1%)
Democratic Republic of the Congo 2013	48696	3.1% (3%,3.3%)	18777	1.1% (1%,1.3%)
Eritrea 1995	11009	4.4% (4.1%,4.8%)	3721	1.7% (1.3%,2.1%)
Ethiopia 2000	29432	4.7% (4.5%,4.9%)	10384	1.4% (1.1%,1.6%)
Ethiopia 2005	30228	3.9% (3.7%,4.2%)	9842	1.1% (0.9%,1.3%)
Ethiopia 2010	34929	2.8% (2.6%,3%)	11802	0.8% (0.6%,0.9%)
Gabon 2000	13354	2.3% (2.1%,2.6%)	4580	0.8% (0.5%,1%)
Gabon 2012	17165	1.9% (1.7%,2.1%)	6446	0.4% (0.3%,0.6%)
Guinea 1999	16473	2.9% (2.7%,3.2%)	5587	1% (0.7%,1.3%)
Guinea 2005	18221	2.8% (2.5%,3%)	6066	0.9% (0.6%,1.1%)
Guinea 2012	21445	3.2% (3%,3.4%)	7137	1% (0.8%,1.3%)
Kenya 1998	17197	2.8% (2.6%,3.1%)	5475	1% (0.7%,1.3%)
Kenya 2003	16291	3.9% (3.6%,4.2%)	5840	1.3% (1%,1.6%)
Kenya 2014	69040	2.8% (2.7%,3%)	21970	0.7% (0.6%,0.8%)
Lesotho 2004	14886	8.3% (7.8%,8.7%)	4560	2.3% (1.9%,2.8%)
Lesotho 2009	16016	9.7% (9.2%,10.1%)	4973	2.2% (1.8%,2.6%)
Lesotho 2014	14166	9.1% (8.7%,9.6%)	4277	2.1% (1.7%,2.5%)
Liberia 2006	16346	2.6% (2.4%,2.9%)	6149	1% (0.8%,1.3%)
Liberia 2013	22513	1.9% (1.8%,2.1%)	8188	0.7% (0.5%,0.9%)
Madagascar 1992	14116	4.3% (3.9%,4.6%)	5212	1.1% (0.9%,1.4%)
Madagascar 1997	15786	3.3% (3%,3.6%)	6055	1.2% (0.9%,1.4%)
Madagascar 2003	16831	2.6% (2.4%,2.9%)	5771	0.8% (0.5%,1%)
Madagascar 2008	39832	2.5% (2.4%,2.7%)	13237	0.7% (0.6%,0.9%)
Malawi 1992	11649	4% (3.7%,4.4%)	3999	1.4% (1.1%,1.8%)
Malawi 2000	29478	5.1% (4.9%,5.4%)	11191	1% (0.8%,1.2%)
Malawi 2004	29294	5.5% (5.2%,5.8%)	10941	1.3% (1.1%,1.6%)
Malawi 2010	58245	4.4% (4.2%,4.5%)	20273	0.9% (0.8%,1%)
Mali 2006	35865	2.1% (1.9%,2.2%)	13502	0.7% (0.6%,0.8%)
Mali 2012	29720	1.6% (1.5%,1.8%)	10852	0.5% (0.4%,0.6%)

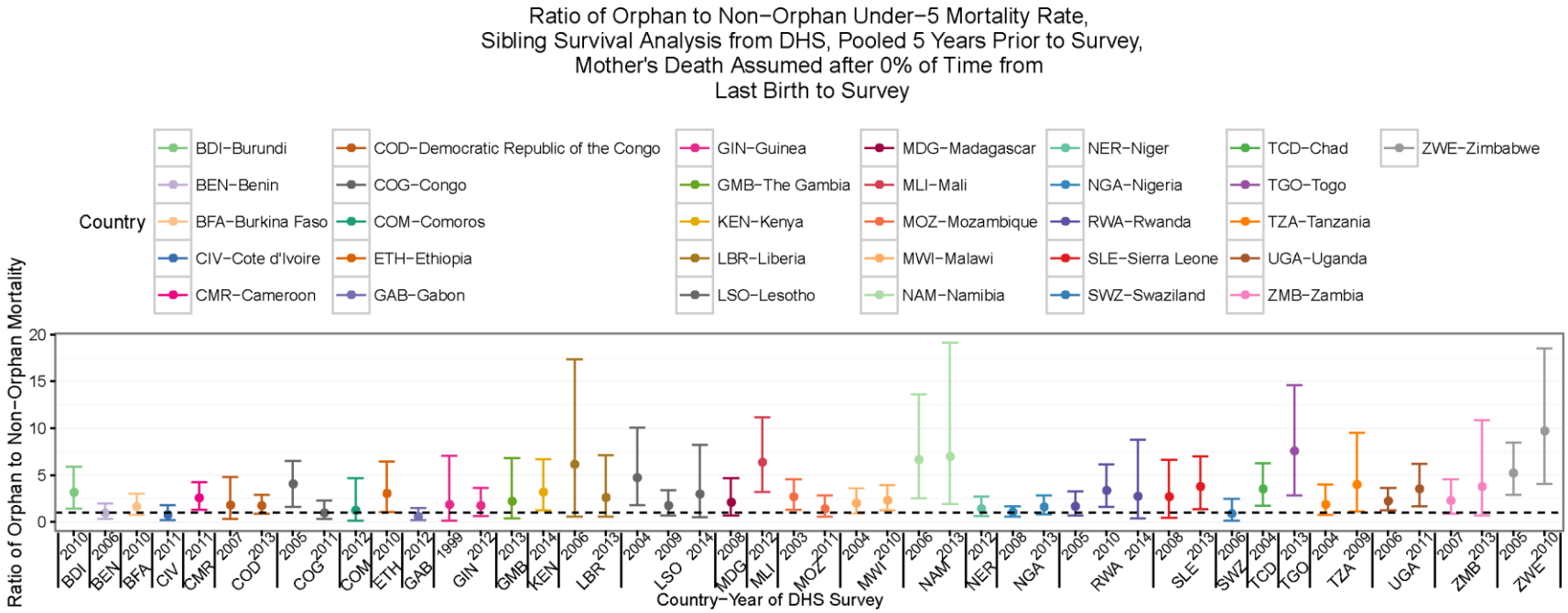
Mauritania 2000	15728	3.2% (2.9%,3.5%)	4795	1.2% (0.9%,1.5%)
Mozambique 1997	20238	5.3% (5%,5.6%)	6941	1.3% (1.1%,1.6%)
Mozambique 2003	29153	3.9% (3.7%,4.1%)	10707	0.9% (0.7%,1.1%)
Mozambique 2011	29822	4.1% (3.9%,4.3%)	10822	1.2% (1%,1.4%)
Namibia 1992	11375	2.1% (1.8%,2.4%)	4279	0.8% (0.6%,1.1%)
Namibia 2000	13014	3.5% (3.2%,3.8%)	4448	1% (0.7%,1.3%)
Namibia 2006	16735	5.9% (5.6%,6.3%)	5746	1.4% (1.1%,1.8%)
Namibia 2013	15857	4.5% (4.2%,4.8%)	5835	1.5% (1.1%,1.8%)
Niger 2006	24943	2.7% (2.5%,2.9%)	9158	1% (0.8%,1.3%)
Niger 2012	35172	2.4% (2.3%,2.6%)	13001	0.9% (0.7%,1%)
Nigeria 2008	70882	1.8% (1.7%,1.9%)	27080	0.7% (0.6%,0.8%)
Nigeria 2013	80356	1.6% (1.5%,1.7%)	30086	0.5% (0.4%,0.6%)
Rwanda 2000	21305	8.8% (8.4%,9.2%)	7406	2% (1.6%,2.3%)
Rwanda 2005	21956	5.9% (5.6%,6.2%)	8233	1.1% (0.9%,1.4%)
Rwanda 2010	24938	3.1% (2.9%,3.3%)	8976	0.6% (0.4%,0.7%)
Rwanda 2014	23495	2% (1.8%,2.2%)	7948	0.5% (0.4%,0.7%)
Sao Tome and Principe 2008	5914	1.1% (0.8%,1.4%)	2133	0.1% (0%,0.3%)
Senegal 1992	15057	2% (1.8%,2.2%)	5666	0.6% (0.4%,0.8%)
Sierra Leone 2008	19829	4% (3.8%,4.3%)	6378	1.7% (1.4%,2%)
Sierra Leone 2013	34344	3.4% (3.2%,3.6%)	12255	1.8% (1.5%,2%)
South Africa 1998	20438	2.5% (2.3%,2.7%)	5727	1.2% (1%,1.6%)
Swaziland 2006	9481	8.3% (7.7%,8.8%)	3136	2.2% (1.8%,2.8%)
Tanzania 1996	18403	2.8% (2.6%,3%)	6612	0.7% (0.5%,0.8%)
Tanzania 2004	22819	2.9% (2.7%,3.1%)	8663	0.7% (0.6%,0.9%)
Tanzania 2009	23122	2.5% (2.2%,2.7%)	8413	0.5% (0.3%,0.6%)
The Gambia 2013	24055	2% (1.8%,2.2%)	9106	0.6% (0.4%,0.7%)
Togo 1998	21226	3% (2.8%,3.2%)	6836	0.8% (0.6%,1%)
Togo 2013	21552	2.5% (2.3%,2.7%)	7235	0.6% (0.4%,0.8%)
Uganda 1995	18035	5.3% (5%,5.6%)	7002	1.6% (1.3%,1.9%)
Uganda 2000	18997	5.6% (5.3%,5.9%)	7182	1.2% (1%,1.5%)
Uganda 2006	23660	5.7% (5.4%,5.9%)	8584	1.4% (1.2%,1.7%)
Uganda 2011	22713	4% (3.7%,4.3%)	8320	0.9% (0.7%,1.1%)
Zambia 1996	18488	4.3% (4%,4.6%)	6807	1.1% (0.8%,1.3%)
Zambia 2001	18174	6% (5.6%,6.3%)	6696	1.3% (1.1%,1.6%)
Zambia 2007	17371	5.6% (5.3%,5.9%)	6459	1% (0.8%,1.3%)
Zambia 2013	40875	3.6% (3.4%,3.8%)	14057	0.9% (0.7%,1%)
Zimbabwe 1994	13876	2.7% (2.4%,3%)	4434	0.8% (0.6%,1.1%)
Zimbabwe 1999	12335	4.8% (4.4%,5.2%)	3909	1.5% (1.1%,1.9%)
Zimbabwe 2005	18742	8.8% (8.4%,9.3%)	5988	2.4% (2%,2.8%)
Zimbabwe 2010	17809	8% (7.6%,8.4%)	6113	1.7% (1.4%,2.1%)

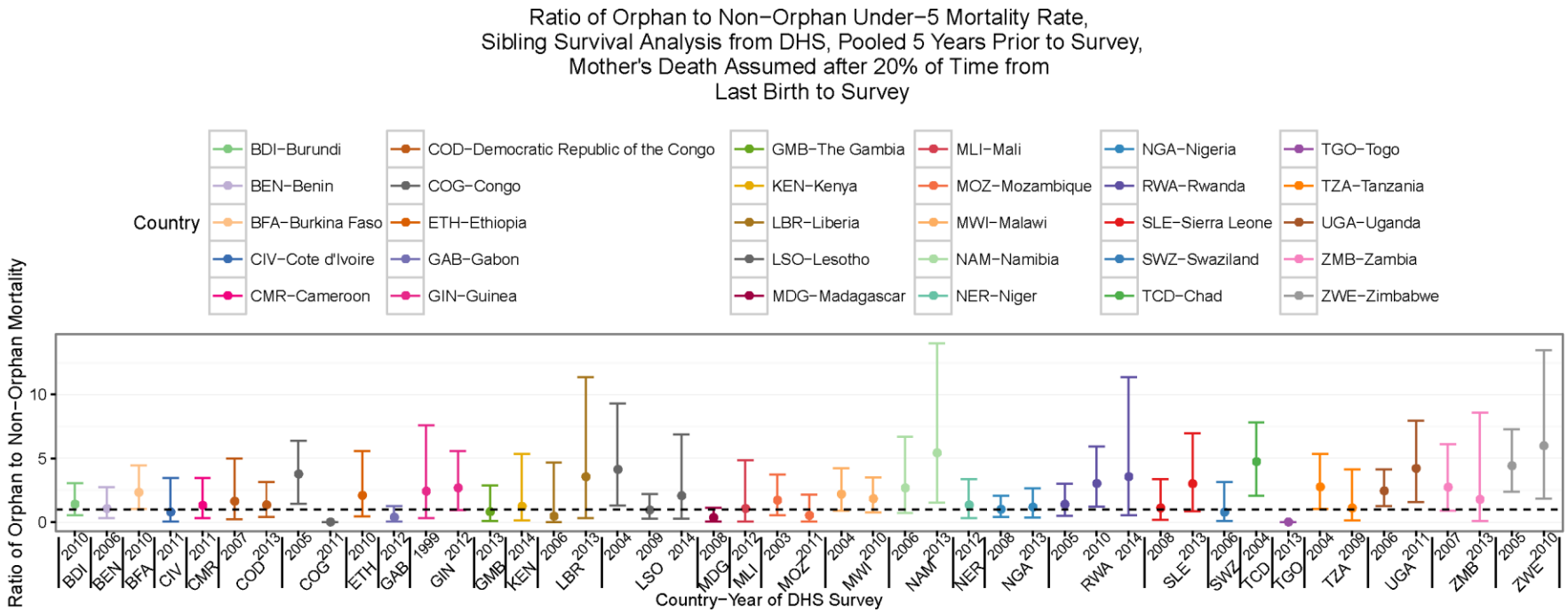
Supplementary Table: Sibling Survival Survey Characteristics								
	Respondents				Person-Years of Sibling Exposure Under -5#		Deaths of Siblings Under-5#	
DHS Survey	Total	Mothers Alive	Mother Deceased^	Mother Survival Unknown	Mother Alive	Mother Deceased*	Mother Alive	Mother Deceased*
Burundi 2010	1431	1204	208	19	4675	216	100	17
Benin 2006	1756	1663	91	2	4927	113	110	2
Burkina Faso 2010	2089	1907	136	46	6341	216	127	7
Cote d'Ivoire 2011	1148	1018	103	27	2661	115	84	5
Cameroon 2011	2139	1873	181	85	5181	271	116	13
Democratic Republic of the Congo 2007	1147	1057	89	1	3462	79	123	9
Democratic Republic of the Congo 2013	2301	2134	164	3	8833	216	283	12
Congo 2005	1301	1191	110	0	3151	75	91	9
Congo 2011	1308	1191	81	36	3836	138	62	2
Comoros 2012	773	722	32	19	1911	67	14	0
Ethiopia 2010	2339	2164	147	28	7508	187	139	9
Gabon 2012	1096	990	69	37	2990	131	62	5
Guinea 1999	264	246	18	0	825	36	21	2
Guinea 2012	1141	1009	109	23	3023	156	82	7
The Gambia 2013	1340	1258	81	1	3954	105	31	2
Kenya 2014	1755	1622	131	2	5166	83	49	0
Liberia 2006	756	725	30	1	2010	23	32	2
Liberia 2013	1140	1079	61	0	2972	95	54	2
Lesotho 2004	1006	852	143	11	1840	147	19	8
Lesotho 2009	1043	819	223	1	1806	185	28	3
Lesotho 2014	948	750	197	1	1389	62	16	3
Madagascar 2008	2440	2161	182	97	6870	352	117	9
Mali 2012	1101	1022	78	1	2464	44	35	3
Mozambique 2003	1430	1297	133	0	4012	178	103	11
Mozambique 2011	1876	1651	223	2	4896	143	74	2
Malawi 2004	1310	1137	172	1	3712	133	113	8
Malawi 2010	3333	2916	372	45	10511	387	203	22
Namibia 2006	1256	1083	173	0	2712	158	40	10
Namibia 2013	1013	892	119	2	2084	92	11	4
Niger 2012	1151	1001	58	92	3579	242	106	12
Nigeria 2008	3932	3527	180	225	10762	614	275	14
Nigeria 2013	4925	4557	243	125	13213	463	302	13
Rwanda 2005	1575	1306	263	6	4282	84	91	4
Rwanda 2010	1866	1619	211	36	5592	160	80	3
Rwanda 2014	1718	1597	119	2	4775	88	49	3
Sierra Leone 2008	629	547	51	31	982	66	37	6
Sierra Leone 2013	2353	2162	189	2	5055	162	123	8
Swaziland 2006	734	609	121	4	1686	116	26	1
Chad 2004	880	800	80	0	3437	103	68	6
Togo 2013	1011	944	67	0	2606	45	39	4
Tanzania 2004	1353	1249	103	1	4776	106	97	3
Tanzania 2009	1417	1285	113	19	4757	124	59	4
Uganda 2006	1210	1056	152	2	4504	142	116	5
Uganda 2011	1225	1087	127	11	4670	189	107	10
Zambia 2007	976	830	146	0	2760	119	52	7
Zambia 2013	2109	1882	224	3	7120	145	103	6
Zimbabwe 2005	2344	1932	407	5	4677	310	53	20
Zimbabwe 2010	1179	914	261	4	1910	109	18	5

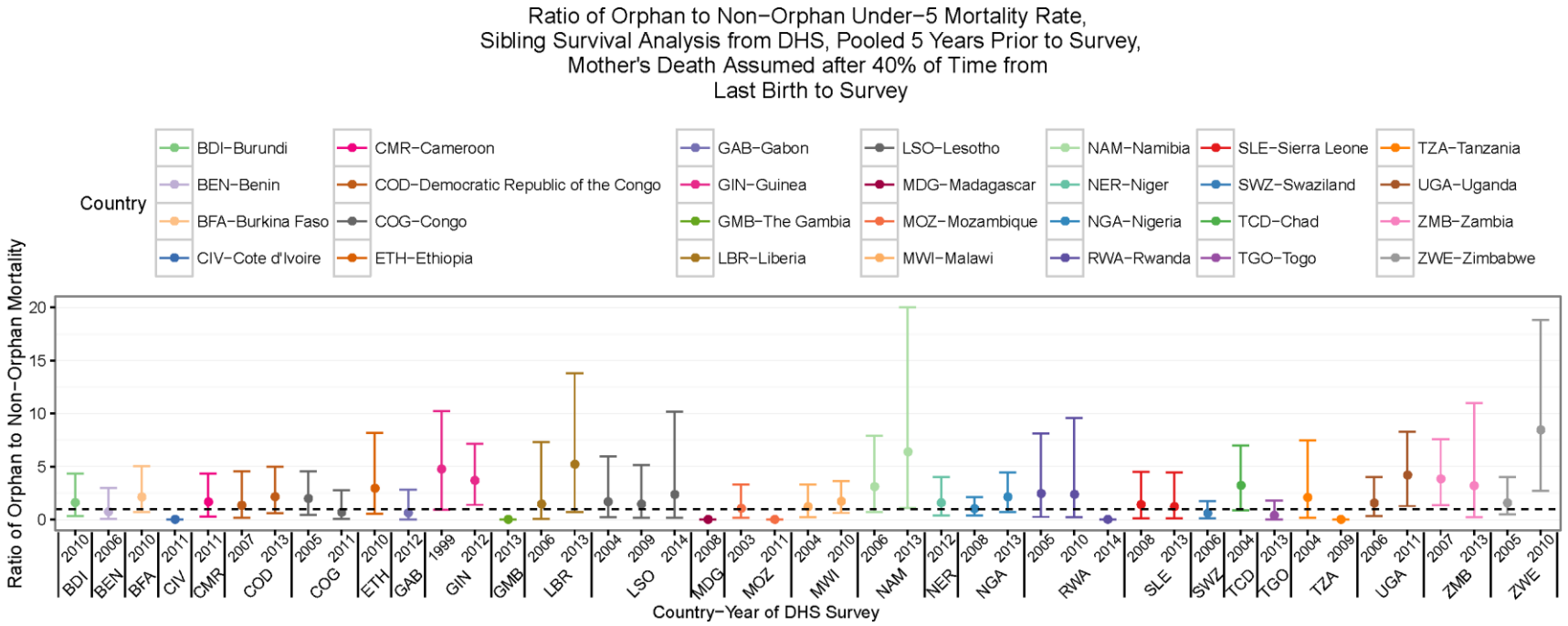
* counts of person-years and deaths to deceased mothers assuming mother died just after birth of last child

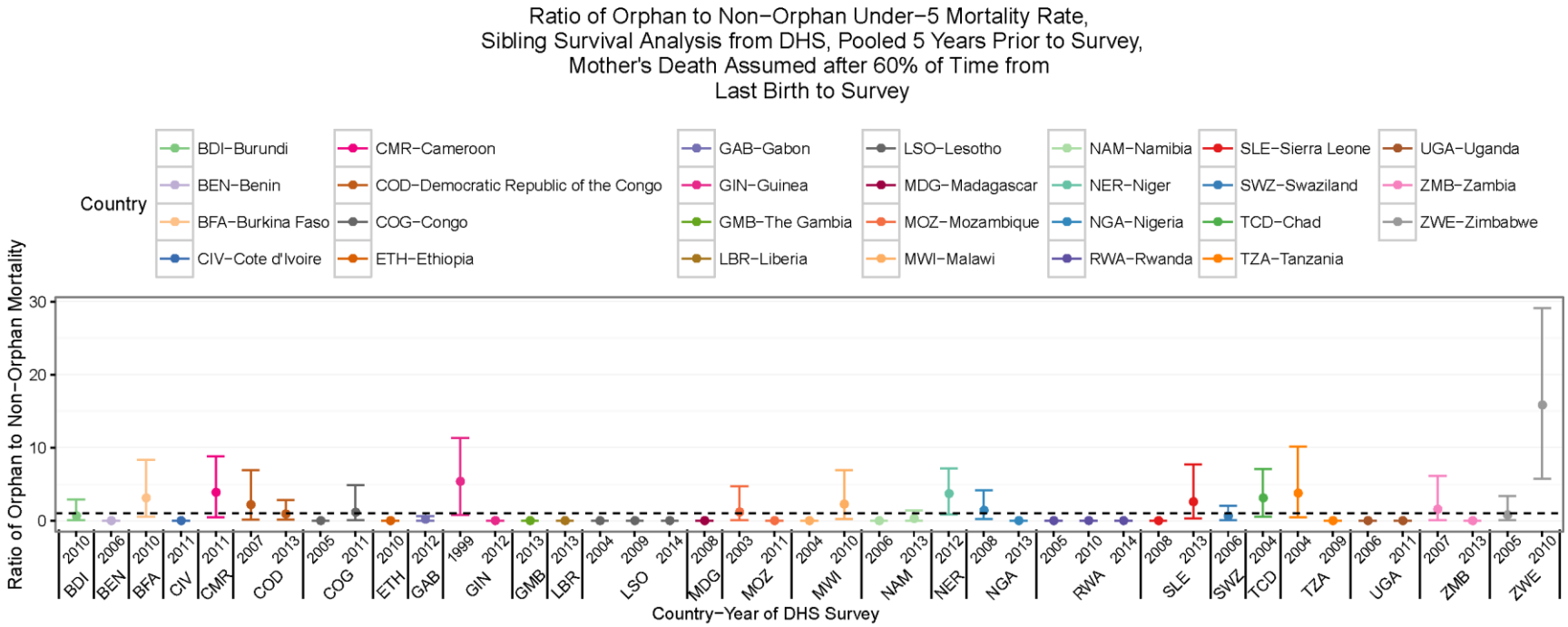
^ mother deceased at time of survey

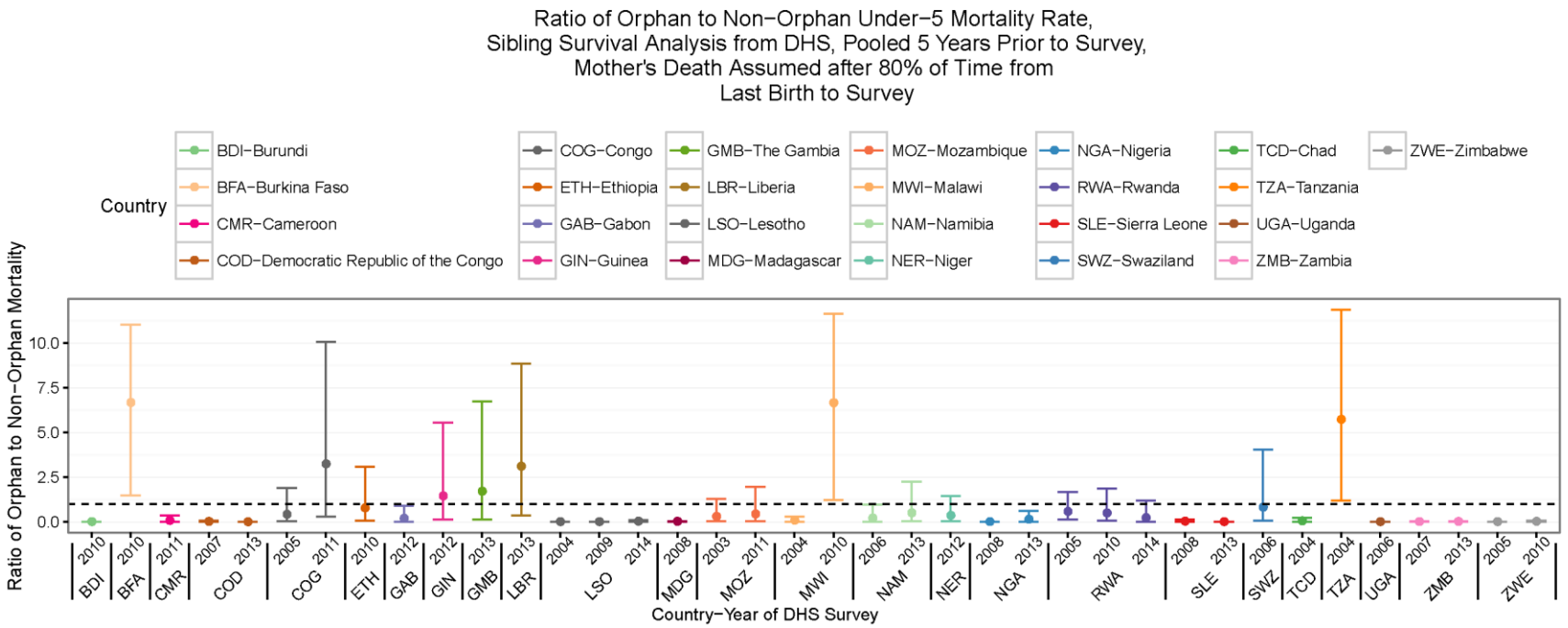
counted in the 5 years before the survey, though estimation process uses data from longer time period


Ratio of Orphan to Non−Orphan Under−5 Mortality Rate,
Sibling Survival Analysis from DHS, Pooled 5 Years Prior to Survey,
Mother's Death Assumed after 0% of Time from
Last Birth to Survey
Country
BDI−Burundi
BEN−Benin
BFA−Burkina Faso
CIV−Cote d'Ivoire
CMR−Cameroon
COD−Democratic Republic of the Congo
COG−Congo
COM−Comoros
ETH−Ethiopia
GAB−Gabon
GIN−Guinea
GMB−The Gambia
KEN−Kenya
LBR−Liberia
LSO−Lesotho
MDG−Madagascar
MLI−Mali
MOZ−Mozambique
MWI−Malawi
NAM−Namibia
NER−Niger
NGA−Nigeria
RWA−Rwanda
SLE−Sierra Leone
SWZ−Swaziland
TCD−Chad
TGO−Togo
TZA−Tanzania
UGA−Uganda
ZMB−Zambia
ZWE−Zimbabwe
Ratio of Orphan to Non−Orphan Mortality
0
5
10
15
20
BDI 2010
BEN 2006
BFA 2010
CIV 2011
CMR 2011
COD 2007
COG 2013
COM 2005
ETH 2011
GAB 2012
GIN 1999
GMB 2012
KEN 2013
LBR 2014
LSO 2006
MDG 2013
MLI 2004
MOZ 2009
MWI 2011
NAM 2010
NER 2012
NGA 2008
RWA 2013
SLE 2005
SWZ 2010
TCD 2014
TGO 2008
TZA 2012
UGA 2004
ZMB 2009
ZWE 2006
2011
2007
2013
2005
2010
Country−Year of DHS Survey


Ratio of Orphan to Non−Orphan Under−5 Mortality Rate,
Sibling Survival Analysis from DHS, Pooled 5 Years Prior to Survey,
Mother's Death Assumed after 20% of Time from
Last Birth to Survey
Country
BDI−Burundi
BEN−Benin
BFA−Burkina Faso
CIV−Cote d'Ivoire
CMR−Cameroon
COD−Democratic Republic of the Congo
COG−Congo
ETH−Ethiopia
GAB−Gabon
GIN−Guinea
GMB−The Gambia
KEN−Kenya
LBR−Liberia
LSO−Lesotho
MDG−Madagascar
MLI−Mali
MOZ−Mozambique
MWI−Malawi
NAM−Namibia
NER−Niger
NGA−Nigeria
RWA−Rwanda
SLE−Sierra Leone
SWZ−Swaziland
TCD−Chad
TGO−Togo
TZA−Tanzania
UGA−Uganda
ZMB−Zambia
ZWE−Zimbabwe
Ratio of Orphan to Non−Orphan Mortality
0
5
10
2010 2006 2010 2011 2011 2007 2013 2005 2011 2010 2012 1999 2012 2013 2014 2006 2013 2004 2009 2014 2008 2012 2003 2011 2004 2010 2006 2013 2012 2008 2013 2005 2010 2014 2008 2013 2006 2004 2013 2004 2009 2006 2011 2007 2013 2005 2010
BDI BEN BFA CIV CMR COD COG ETH GAB GIN GMB KEN LBR LSO MDG MLI MOZ MWI NAM NER NGA RWA SLE SWZ TCD TGO TZA UGA ZMB ZWE
Country−Year of DHS Survey


Ratio of Orphan to Non−Orphan Under−5 Mortality Rate,
Sibling Survival Analysis from DHS, Pooled 5 Years Prior to Survey,
Mother's Death Assumed after 40% of Time from
Last Birth to Survey
Country
BDI−Burundi
BEN−Benin
BFA−Burkina Faso
CIV−Cote d'Ivoire
CMR−Cameroon
COD−Democratic Republic of the Congo
COG−Congo
ETH−Ethiopia
GAB−Gabon
GIN−Guinea
GMB−The Gambia
LBR−Liberia
LSO−Lesotho
MDG−Madagascar
MOZ−Mozambique
MWI−Malawi
NAM−Namibia
NER−Niger
NGA−Nigeria
RWA−Rwanda
SLE−Sierra Leone
SWZ−Swaziland
TCD−Chad
TGO−Togo
TZA−Tanzania
UGA−Uganda
ZMB−Zambia
ZWE−Zimbabwe
Ratio of Orphan to Non−Orphan Mortality
0
5
10
15
20
2010 2006 2010 2011 2011 2007 2013 2005 2011 2010 2012 1999 2012 2013 2006 2013 2004 2009 2014 2008 2003 2011 2004 2010 2006 2013 2012 2008 2013 2005 2010 2014 2008 2013 2006 2004 2013 2004 2009 2006 2011 2007 2013 2005 2010
BDI BEN BFA CIV CMR COD COG ETH GAB GIN GMB LBR LSO MDG MOZ MWI NAM NER NGA RWA SLE SWZ TCD TGO TZA UGA ZMB ZWE
Country−Year of DHS Survey


Ratio of Orphan to Non−Orphan Under−5 Mortality Rate,
Sibling Survival Analysis from DHS, Pooled 5 Years Prior to Survey,
Mother's Death Assumed after 60% of Time from
Last Birth to Survey
Country
BDI−Burundi
BEN−Benin
BFA−Burkina Faso
CIV−Cote d'Ivoire
CMR−Cameroon
COD−Democratic Republic of the Congo
COG−Congo
ETH−Ethiopia
GAB−Gabon
GIN−Guinea
GMB−The Gambia
LBR−Liberia
LSO−Lesotho
MDG−Madagascar
MOZ−Mozambique
MWI−Malawi
NAM−Namibia
NER−Niger
NGA−Nigeria
RWA−Rwanda
SLE−Sierra Leone
SWZ−Swaziland
TCD−Chad
TZA−Tanzania
UGA−Uganda
ZMB−Zambia
ZWE−Zimbabwe
Ratio of Orphan to Non−Orphan Mortality
30
20
10
0
2010 2006 2010 2011 2011 2007 2013 2005 2011 2010 2012 1999 2012 2013 2013 2004 2009 2014 2008 2003 2011 2004 2010 2006 2013 2012 2008 2013 2005 2010 2014 2008 2013 2006 2004 2004 2009 2006 2011 2007 2013 2005 2010
BDI BEN BFA CIV CMR COD COG ETH GAB GIN GMB LBR LSO MDG MOZ MWI NAM NER NGA RWA SLE SWZ TCD TZA UGA ZMB ZWE
Country−Year of DHS Survey


Ratio of Orphan to Non−Orphan Under−5 Mortality Rate,
Sibling Survival Analysis from DHS, Pooled 5 Years Prior to Survey,
Mother's Death Assumed after 80% of Time from
Last Birth to Survey
Country
BDI−Burundi
BFA−Burkina Faso
CMR−Cameroon
COD−Democratic Republic of the Congo
COG−Congo
ETH−Ethiopia
GAB−Gabon
GIN−Guinea
GMB−The Gambia
LBR−Liberia
LSO−Lesotho
MDG−Madagascar
MOZ−Mozambique
MWI−Malawi
NAM−Namibia
NER−Niger
NGA−Nigeria
RWA−Rwanda
SLE−Sierra Leone
SWZ−Swaziland
TCD−Chad
TZA−Tanzania
UGA−Uganda
ZMB−Zambia
ZWE−Zimbabwe
Ratio of Orphan to Non−Orphan Mortality
10.0
7.5
5.0
2.5
0.0
2010 BDI
2010 BFA
2011 CMR
2007 COD
2013
2005
2011 COG
2010 ETH
2012 GAB
2012 GIN
2013 GMB
2013 LBR
2004
2009
2014 LSO
2008 MDG
2003
2011 MOZ
2004
2010 MWI
2006
2013 NAM
2012 NER
2008
2013 NGA
2005
2010 RWA
2014
2008
2013 SLE
2006 SWZ
2004 TCD
2004 TZA
2006 UGA
2007
2013 ZMB
2005
2010 ZWE
Country−Year of DHS Survey