USAID
FROM THE AMERICAN PEOPLE

DHS WORKING PAPERS

Accumulating Birth Histories across Surveys for Improved Estimates of Child Mortality

Laura Schmidt
Mahmoud Elkasabi

2020 No. 177

DEMOGRAPHIC AND HEALTH SURVEYS

November 2020

DHS Working Papers No. 177

Accumulating Birth Histories across Surveys for Improved Estimates of Child Mortality

Laura Schmidt[1]
Mahmoud Elkasabi[2]

ICF
Rockville, Maryland, USA

November 2020

[1] Johannes Gutenberg University Mainz, Mainz, Germany
[2] The DHS Program, ICF

Corresponding author: Mahmoud Elkasabi, The DHS Program, ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850, USA; phone: +1 301-572-0484; fax: +1 301-407-6501; email: Mahmoud.Elkasabi@icf.com

Editor: Diane Stoy
Document Production: Joan Wardell

This study was conducted with support from the United States Agency for International Development (USAID) through The DHS Program (#720-OAA-18C-00083). The views expressed are those of the authors and do not necessarily reflect the views of USAID or the United States Government.

The DHS Program assists countries worldwide in the collection and use of data to monitor and evaluate population, health, and nutrition programs. Additional information about The DHS Program can be obtained from ICF, 530 Gaither Road, Suite 500, Rockville, MD 20850 USA; telephone: +1 301-572-0200, fax: +1 301-572-0999, email: info@DHSprogram.com, Internet: www.DHSprogram.com.

Recommended citation:

Schmidt, Laura, and Mahmoud Elkasabi. 2020. *Accumulating Birth Histories across Surveys for Improved Estimates of Child Mortality*. DHS Working Papers No. 177. Rockville, Maryland, USA: ICF.

CONTENTS

TABLES AND FIGURES

ABSTRACT

This paper outlines a framework to guide the process of accumulating birth histories across different DHS/MICS surveys. We focus on three rates that are essential to monitoring progress in meeting the UN SDGs: the NNMR, IMR, and U5MR. Based on DHS survey data from Senegal and DHS and MICS survey data from Malawi, we show that the framework for accumulating birth histories across different surveys reduces the width of confidence intervals around the estimated mortality rates. The estimated standard errors of the cumulated rates are about 50-60% less than their counterparts from separate surveys. In addition, accumulating birth histories smooths fluctuations in time series for national and subnational mortality rates. It establishes time trends that are more stable and reliable than those based on data from single survey data. The results of this paper stress the importance of cumulating survey data to identifying reliable trends and correctly evaluating policies designed to reduce mortality rates.

Key words: Childhood mortality rates, neonatal mortality rate, infant mortality rate, under-5 mortality rate, cumulating survey data, DHS surveys, MICS surveys

1 INTRODUCTION

Since the mid-1980s, the Demographic and Health Surveys (DHS) Program has conducted hundreds of national surveys that measure health and demographic indicators. Measuring mortality rates among children has been one of the main objectives of DHS surveys since the beginning. Such rates are calculated with data from birth history reported by interviewed women, especially date of birth and survival status of each live birth, and age at death of each deceased child. The childhood mortality indicators that are calculated and released in the final report of each DHS survey include the neonatal mortality rate (NNMR), post-neonatal mortality rate (PNMR), infant mortality rate (IMR), child mortality rate (CMR), and the under-5 mortality rate (U5MR). Three of the childhood mortality indicators (NNMR, IMR, and U5MR) are essential to monitoring progress in meeting the United Nations (UN) Sustainable Development Goals (SDGs) (Abel et al. 2016; United Nations Inter-agency Group for Child Mortality Estimation 2018).

Reliable estimates of childhood mortality rates are crucial to evaluating progress in achieving the SDG goals. This includes estimates within reasonable confidence intervals (CI), either at the national or subnational levels. In the DHS final reports, national rates are calculated on birth histories of 5 years before the survey, while regional rates are calculated on birth histories of 10 years before the survey. The longer reference period of the regional rates allows for more birth history data to contribute to the regional rates, which results in lower standard errors (SE) and more narrow CIs around the rates.

During the past 35 years, many countries collected more than 5 DHS surveys, which were typically 5 years apart. This includes, but is not limited to, Bangladesh, Colombia, Dominican Republic, Egypt, Ghana, Indonesia, Jordan, Kenya, Mali, Nigeria, Philippines, Senegal, Tanzania, Uganda, Zambia, and Zimbabwe (https://dhsprogram.com/). Such long series of surveys allow researchers to study the change in survey estimates across years. Traditionally, the DHS surveys have been designed and used for measuring periodic changes across surveys in addition to "current" estimates. The development of DHS surveys concentrated on tracking survey estimates across surveys within and between countries, and has encouraged the development of a harmonized set of survey tools and procedures for all survey stages. The harmonized toolkit includes standard sampling designs, questionnaires, protocols for anthropometric measurements and blood testing, as well as algorithms for survey estimates and tabulation plans. Most of the tools available on the DHS Program website (https://dhsprogram.com/) have been used extensively to collect surveys other than the DHS surveys. For example, the DHS program produced hundreds of *comparative reports* in which many indicators were tracked across years or between countries. Such harmonization allows for the combination of DHS surveys with other national surveys that collect similar data, such as the UNICEF's Multiple Indicator Cluster Survey (MICS).

Although DHS surveys are cross-sectional surveys that produce estimates based on the "current" situation, there are some data, such as birth histories, that are collected based on reports of retrospective events or actions. The birth histories that are collected from each interviewed woman include sex and date of birth of each child, and age at death of each deceased child. This information allows for fertility rates and childhood mortality rates to be calculated from the DHS survey data. Although the calculation of DHS rates is based on a defined reference period, in which only events that happened during the reference period contribute to the calculation of the rate, full birth histories are collected from all interviewed women. Such full birth histories allow for calculating the rates based on calendar years and reference periods beyond the period in the DHS final reports. For example, with Zambia DHS 2018, in addition to the published national childhood mortality rates that are based on a reference period of 5 years before the survey (5 years that ends with the date of interview), rates can also be calculated based on other reference periods, such as December 2013 to January 2009 (Zambia Statistics Agency et al. 2019).

Given the overlap in birth histories across successive surveys, different surveys collect data about birth histories in the same calendar years. For example, data about births and deceased children during 2003 were collected in both Malawi DHS 2015-16 and Malawi DHS 2010. Unfortunately, such overlap is not fully utilized. Births histories of 2003 collected in Malawi DHS 2010 contribute to the regional childhood mortality rates calculated on the 10 years before the survey. In contrast, birth histories of 2003 collected in Malawi DHS 2015-16 do not contribute to any of the childhood mortality rates released in the DHS final report. Moreover, data about birth histories can be collected in other reliable surveys, such as the MICS, and used to produce survey-level indicators. For example, data about births and deceased children before 2014 are collected in both Malawi MICS 2013-14 and Malawi DHS 2015-16 (National Statistical Office and ICF 2017; National Statistical Office 2015; National Statistical Office and ICF Macro 2011). Such overlap in birth histories across surveys, within DHS surveys, or between DHS and MICS surveys, can be exploited for the benefit of childhood mortality indicators. Other than Pullum and Assaf (2016), who pooled DHS data for 16 countries and re-calculated fertility rates by aggregating births and exposure across surveys, we have not found any studies that utilize the overlapping birth histories.

In general, the literature lacks a framework that guides the process of accumulating birth histories across several national surveys, including DHS surveys and other surveys such as MICS surveys. Birth history data from consecutive DHS surveys can be cumulated in one pooled dataset so that all reported birth histories of a specific calendar year "*A*" are cumulated in one dataset. Such a cumulated dataset can be used for different purposes such as producing 1) national mortality rates with better precision; 2) mortality rates based on shorter reference periods and 3) regional mortality rates with better precision and/or based on shorter reference periods. In this paper, we outline a framework to guide the process of accumulating birth histories across different DHS/MICS surveys. Section 2 is a brief background about the data cumulation techniques, the DHS birth histories, and the childhood mortality rates. In Section 3, we outline a framework for accumulating birth history data using data from DHS surveys. After describing the accumulation framework, we have illustrated and experimented with several surveys in Section 4. Finally, we conclude the paper with a discussion in Section 5.

2 BACKGROUND

2.1 Techniques for Cumulating Periodic Surveys

Cumulating data across several surveys achieves many goals, such as 1) increasing sample size, which produces reliable estimates with reduced estimated variance on the national level, or on small geographic and non-geographic domains; 2) estimating change across different surveys (Kalton 2009; Robert and Binder 2009; Thomas and Wannell 2009). The cumulating technique has also been used to combine different samples in multiple frame designs to overcome frame non-coverage problems (Elkasabi 2015; Elkasabi et al. 2015; Lohr 2011). The technique has also has been used to combine rolling samples of the same survey. For example, the Senegal DHS 2012-14 is based on a cumulated sample from the Senegal DHS 2012-13 and the Senegal DHS 2014 (Agence Nationale de la Statistique et de la Démographie and ICF International 2015).

Kish (1999) defined "cumulating periodic surveys" as the process of cumulating estimates across periodic surveys. Unlike the traditional use of periodic surveys to produce "current" estimates or to measure periodic changes, surveys are combined to produce cumulated estimates. Two broad approaches can be used to cumulate data across surveys: 1) The separate approach, in which the cumulating is done at the estimates level, where periodic estimates are cumulated across surveys; 2) The pooled approach, in which micro data are cumulated across surveys. With the separate approach, periodic estimates are produced separately from each survey, and then combined in a composite estimator. With the pooled approach, micro data from different surveys are cumulated in one "pooled" dataset, and estimates are calculated based on that pooled dataset (Thomas and Wannell 2009).

2.1.1 The separate approach

In the separate approach, cumulated estimates can be calculated as a simple or weighted average of periodic estimates. Let Y denote a population parameter and $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_k$ denote unbiased estimates of Y from k periodic surveys. A cumulated estimate $\hat{y}_c$ can be calculated as:

$$\hat{y}_c = \sum_{i=1}^{k} \hat{\theta}_i \hat{y}_i \quad (1)$$

where $\hat{\theta}_i$ is a cumulating composite factor and $\sum_{i=1}^{k} \hat{\theta}_i = 1$. The cumulated estimate in (1) can also be calculated as a simple average $\hat{y}_c = \sum_{i=1}^{k} \hat{y}_i / k$ where $\hat{\theta}_i = 1/k$. Where samples of periodic surveys are independent, the variance of the cumulated estimate can be calculated as:

$$var(\hat{y}_c) = \sum_{i=1}^{k} \hat{\theta}_i^2 var(\hat{y}_i) \quad (2)$$

Different methods can be used to choose the values of the composite factor $\hat{\theta}_i$, to minimize the variance of the cumulated estimate (2) or to account for the effective sample size for each of the periodic samples. More details about these methods can be found in Chu et al. 1999; Korn and Graubard 1999; and Roberts and Binder 2009.

The main advantage of the separate approach is that it can be easily used to calculate cumulated estimates for simple indicators, such as proportions, means and totals, without having to involve the survey micro data. The approach might be appealing to users who rely on published tables to calculate cumulated estimates. However, this approach is not appropriate for complicated indicators such as the childhood mortality rates identified in this paper. The separate approach does not also allow for the production of new indicators, such as calculating childhood mortality rates based on different reference periods. Finally, the lack of published estimators or their confidence intervals might be a challenge in producing cumulated estimates with estimated confidence intervals.

2.1.2 The pooled approach

Unlike the separate approach, in the pooled approach, individual records, and not estimates, are cumulated across different surveys. Pooled/cumulated estimates are produced by using the same techniques appropriate for single samples. Such techniques might use the same survey weight of each individual record, or a modified version of the weight. Let y_{ij} denote the individual record of y for individual j of survey i and let w_{ij} denote the survey weight for the same individual. A cumulated estimate $\hat{y}_c$ can be calculated as:

$$\hat{y}_c = \sum_i \sum_{j \in i} \theta_{ij} w_{ij} y_{ij} \quad (3)$$

where θ_{ij} is a cumulating adjustment factor that is fixed as $\theta_{ij} = \theta_i$ across all individual records of survey j, with $\sum_{i=1}^{k} \theta_i = 1$. Note that the pooled cumulated estimator in (3) is equivalent to the separate cumulated estimator in (1). In the case of fixed adjustment factors across all surveys, $(\theta_i = 1/k)$, the pooled cumulated estimator in (3) will be equivalent to the simple average version of the separate cumulated estimator.

With data cumulated at the micro-data level, the pooled/cumulated dataset can be analyzed as a single sample from a population, where all single-sample techniques of data analysis are relevant. This allows for the estimation of complicated indicators other than proportions, means, or totals, such as childhood mortality rates. However, this requires access to the micro-data of the separate surveys by data users who can manipulate datasets in order to properly cumulate several surveys.

2.2 Birth History and Childhood Mortality Rates

2.2.1 Birth histories

Birth histories collected in household surveys are a key component in calculating fertility and childhood mortality indicators. The complexity of the collected data varies according to the survey size and objectives. Data about live births might be collected in full, truncated, or summary birth histories. The data might also be collected among other data about all pregnancies in a full pregnancy history (Hill 2013). In most of the DHS surveys, a *full birth history* is collected from all women age 15-49. In some DHS surveys, such as in Egypt, Jordan, Bangladesh, and Afghanistan, full birth histories are only collected from ever-married women age 15–49, because only ever-married women are eligible for the women's interview. In the DHS birth histories, each interviewed woman reports data about all her live-born births, starting from the first to the most recent birth. For each birth, the woman reports the sex, month, and year of birth, current survival status (whether the child is still alive), and current age of surviving children in completed years. For deceased children, the mother reports the age at death, detailed in completed years and months if the child died between 28 days and 24 months, or days if the child died before 28 days (Croft at al. 2018).

Unlike full birth histories, there are partial birth histories, such as the *truncated birth history* and the *summary birth history*. In the truncated histories, data are collected on births up to a particular date before the survey, or an upper limit for number of births for each woman. This reduces fieldwork costs and interviewee fatigue. The reduction in the collected data might result in a sample of births with characteristics that are drastically different from the ones collected with a full birth history. Similarly in the summary histories, only aggregate numbers of children ever-born and still alive are collected. The summary birth histories are often collected in census forms. Unlike the other birth histories, the summary birth history does not allow for direct calculation of standard mortality indicators, such as occurrence-exposure rates. Indirect estimation methods can be used instead (Hill 2013; United Nations 2011).

Unlike birth histories, *pregnancy history* data are collected from interviewed women about all their pregnancies, including pregnancies that resulted in a live birth or pregnancies that did not, such as stillbirths, miscarriages, and abortions. Pregnancy histories are typically collected in a truncated form that covers only births or pregnancies occurring during a fixed period of time, such as the 5 years before the survey (United Nations 2011). In DHS surveys, full pregnancy histories have been collected consistently in Armenia, Azerbaijan, Kazakhstan, Kyrgyz Republic, Moldova, Tajikistan, and Ukraine (MacQuarrie et al. 2018).

2.2.2 Childhood mortality rates

Collecting retrospective birth history in household surveys allows for a direct estimation of childhood mortality indicators. The direct estimation method is commonly used in the DHS and MICS surveys, where the childhood mortality is calculated as occurrences/exposure rates, in which the numerator is the number of deaths and the denominator is the number of births within a reference period. In DHS surveys, five childhood mortality rates are calculated as probabilities of dying during specific age: 1) NNMR: the probability of dying between birth and exact age of 1 month; 2) PNMR: the probability of dying between exact ages 1 month and 1 year, which is usually calculated as the difference between IMR and NNMR; 3) IMR: the probability of dying between birth and exact age of 1 year; 4) CMR: the probability of dying between exact ages 1 and 5 years; and 5) U5MR: the probability of dying between birth and the exact age of 5 years. These probabilities are calculated based on a synthetic cohort life table approach, which is explained in detail in the literature (Croft et al. 2018; Elkasabi 2019; United Nations 2011).

The DHS tradition has been to not report the mortality rates based on calendar years, but based on a window of time that ends with the date of interview. The justification for this practice is that the number of births in the calendar year of interview can be small and statistically unstable if the interviews are done early in the year. Similarly, to avoid statistically unstable single-year rates, the DHS reports the national and regional mortality rates based on a reference period of 5 and 10 years before the survey respectively (Pullum and Assaf 2016). For example, "the past 5 years" or "0-4 [completed] years before the survey" means that the window is the 60 months before the month of interview. For a woman interviewed in June 2019, the time interval for national mortality rates will be June 2014 through May 2019, inclusive, and the time interval for regional mortality rates will be June 2009 through May 2019, inclusive.

3 DATA AND METHODS

3.1 Data

The selection of countries for this analysis was guided by two criteria 1) completion of several DHS surveys during the past 20 years (2000 to 2020) and 2) completion of DHS surveys and MICS surveys during the study interval. Many countries met one of the two criteria, such as Egypt, Ethiopia, Malawi, Nigeria, Senegal, Zambia, and several other countries. Two countries were selected for the analysis in this paper: Senegal and Malawi. In Senegal, due to the large number of DHS surveys collected during the past 10 years, we focused in this study on DHS surveys collected after 2010. As indicated in Table 1, six DHS surveys were completed in Senegal between 2010 and 2017, and four DHS surveys and two MICS surveys were completed in Malawi between 2000 and 2015. For each DHS survey, the births dataset (BR) was used in the data analysis, whereas for the Malawi MICS survey, the birth history (bh) dataset was used in the data analysis. All analysis datasets can be downloaded from the DHS and MICS webpages (https://dhsprogram.com/; https://mics.unicef.org/surveys).

Table 1 The survey and number of women age 15-49 (n) during the study interval (2000-2017 for Senegal; 2000-2015 for Malawi)

Country	Survey (n)	Survey (n)	Survey (n)	Survey (n)	Survey (n)	Survey (n)
Senegal	DHS 2010-11 (15,688)	DHS 2012-13 (8,636)	DHS 2014 (8,488)	DHS 2015 (8,851)	DHS 2016 (8,865)	DHS 2017 (16,787)
Malawi	DHS 2000 (13,020)	DHS 2004 (11,698)	MICS 2006 (26,259)	DHS 2010 (23,020)	MICS 2013 (24,230)	DHS 2015 (24,562)

The birth history data from the analysis datasets overlap as indicated in Figure 1 for Senegal. All six surveys collected data about children deaths and exposure from 2000 through 2011. More specifically, as indicated in Figure 2, all surveys share births and deaths that occurred between 1982 and 2011. For example, childhood mortality rates are typically calculated from the DHS 2010-11 based on a 5-year reference period, highlighted in yellow in Figure 1. With deaths and exposure data collected from the other surveys based on the same reference period, separate childhood mortality rates can be calculated with one set from each survey based on the same reference period as the DHS 2010-11. Birth history data, highlighted in yellow, can be cumulated with the birth history from DHS 2010-11 and cumulated rates can be produced from the cumulated dataset. Deaths and exposures in years after 2011 were collected in surveys fielded after that year. Deaths and exposure in 2017 were only collected in the DHS 2017.

Figure 1 **Overlap for 5-year reference periods by DHS surveys in Senegal**

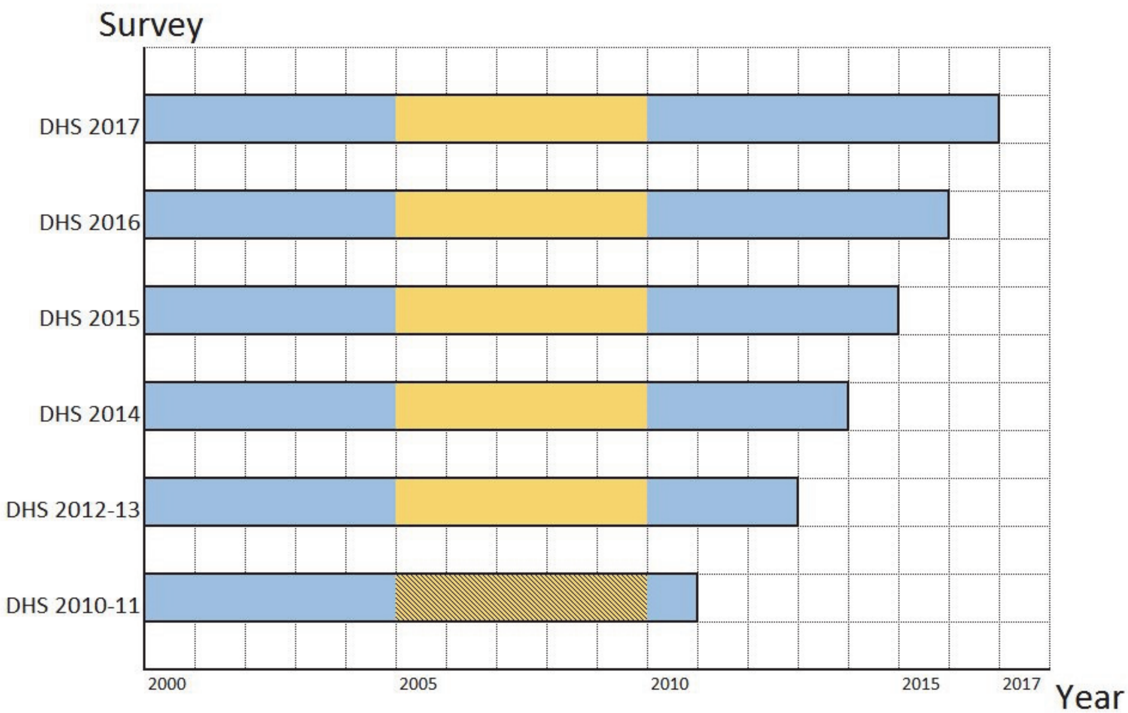
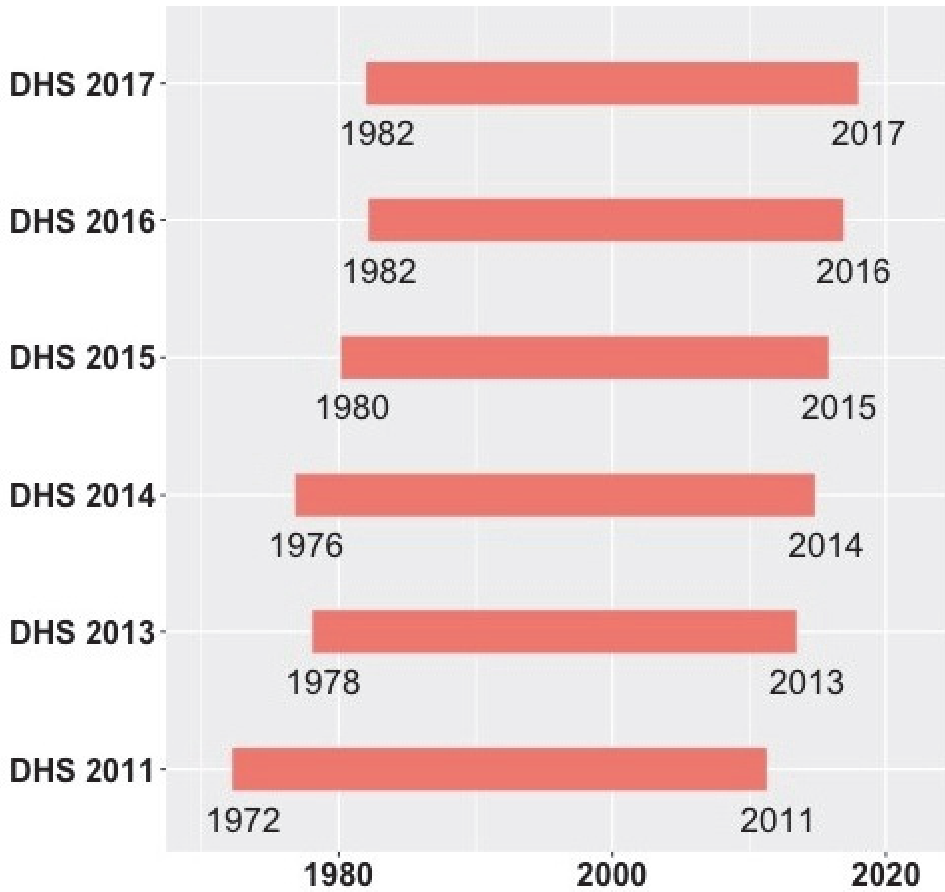


Figure 2 Ranges of dates of birth and death by DHS surveys in Senegal



3.2 Methods

In this analysis, trends of three childhood mortality indicators (NNMR, IMR and U5MR) were tracked in the time interval from 2000 to 2016. The three mortality rates were calculated with standard DHS procedures described in Croft et al. (2018) and Elkasabi (2019). Rates were calculated from separate, one for each survey, and cumulated datasets. In each country, one cumulated dataset was formed by pooling the separate datasets as follows:

1. All separate datasets were pooled in one cumulated dataset.

2. A survey ID variable was constructed to identify each survey within the cumulated dataset.

3. Unique cluster and strata codes were constructed so that clusters and strata from each survey are unique across surveys within the cumulated dataset.

4. A modified survey weight was created by multiplying the survey weight (v005) from each survey by an un-normalization factor that was calculated as the approximated number of women age 15-49 in the population (https://data.worldbank.org/), during the time of the survey, divided by the total number of women age 15-49 interviewed in the survey.

5. The unique cluster, strata, and un-normalized weight were used for the analysis in this paper.

It is worth noting that calculating the childhood mortality rates from the cumulated datasets by using the un-normalized weight is equivalent to the pooled cumulated estimator in (3) with $\theta_i = 1/k$, where k is the number of pooled surveys.

Unlike the rates published in the DHS reports, rates are calculated in this analysis based on calendar years. For example, 5-year mortality rates for 2016 are based on deaths and exposure from January 2012 through December 2016 inclusive. The *chmort* function from the *DHS.rates* R package was used to calculate all rates and their standard errors in this analysis (Elkasabi 2019, 2020).

4 RESULTS

4.1 Childhood Mortality Rates: National Level

Figure 3 reports the national NNMR, IMR, and U5MR of Senegal calculated with separate DHS surveys and cumulated survey data. The separate rates, in the upper panel of the figure, are the rates released in the DHS survey reports of DHS 2010-11, 2012-13, 2014, 2015, 2016, and 2017. From the cumulated data, annual rates based on reference periods of 5 calendar years and one calendar year were calculated during the study period (2000-2016), and presented in the bottom two panels of the figure. In comparison to the DHS released rates, cumulated birth histories across surveys allow for longer time series of childhood mortality trends, such as 2000-2016 or an even longer period due to the complete birth histories collected in the DHS surveys, and trends of annual rates based on reference periods of 5 and 1 calendar years. In addition, with such a large amount of birth history data, the cumulation allows for smoothing fluctuations in time trends. For example, the DHS trends indicate a decrease in childhood mortality rates between 2013 and 2014 and an increase between 2014 and 2015. The cumulated trend based on a 5-year reference period smoothed such fluctuations and allowed for more narrow confidence intervals around the trend curve. As Table 2 indicates, cumulated rates are accompanied by standard errors that are an estimated 60% less than their counterparts from separate surveys. In addition, Figure 3 reveals that cumulating birth histories across surveys allows for the calculation of mortality rates based on shorter reference periods, such as one calendar year, with acceptable confidence intervals, especially for the U5MR.

Table 2 Average standard errors of mortality rates by survey in Senegal

	5 year reference period			*1 year reference period*		
	NNMR	**IMR**	**U5MR**	**NNMR**	**IMR**	**U5MR**
DHS 2011	2.43	3.15	4.70	4.69	5.80	6.99
DHS 2013	3.36	4.42	6.16	6.63	8.25	9.64
DHS 2014	2.98	4.39	6.89	5.24	7.24	9.49
DHS 2015	3.27	4.30	6.23	6.23	7.85	9.44
DHS 2016	3.75	4.66	6.29	7.00	8.46	9.77
DHS 2017	2.34	3.21	4.54	4.71	5.86	6.91
Survey average	3.02	4.02	5.80	5.75	7.24	8.71
Cumulated	1.20	1.57	2.20	2.47	3.03	3.52
Reduction in %	60.27	61.04	62.15	57.08	58.14	59.60

Figure 3 Trends of childhood mortality rates in Senegal based on separate surveys and cumulated data

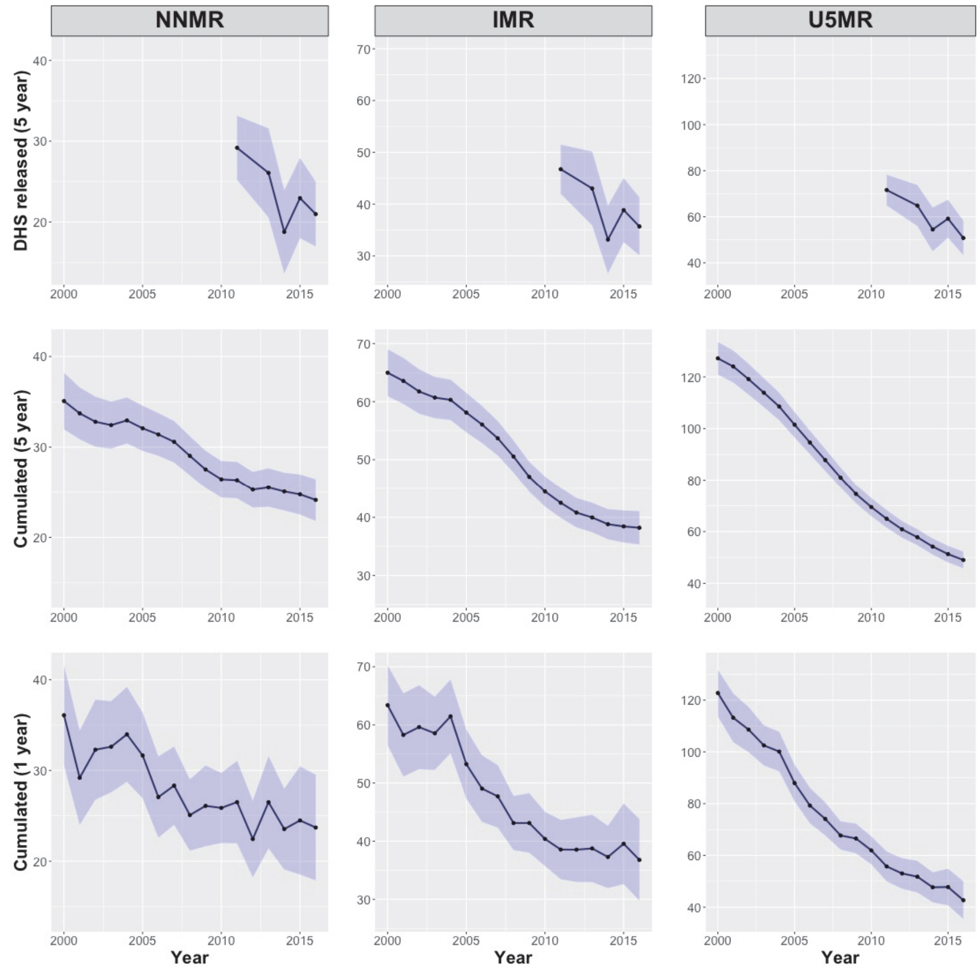
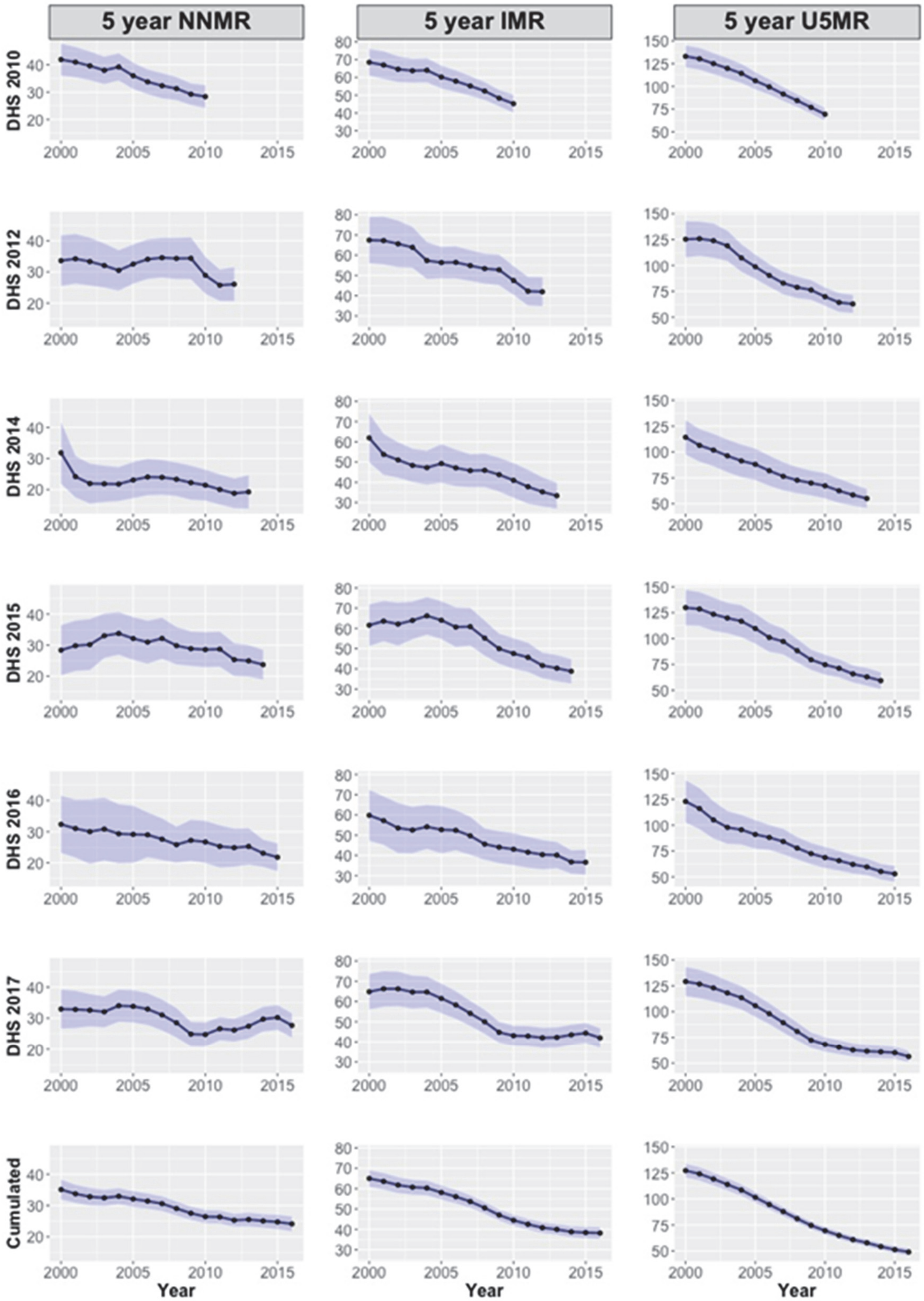


Figure 4 reports trends of annual childhood mortality rates calculated on a reference period of 5 calendar years. Separate trends were calculated based on separate survey datasets and a cumulated dataset. Trends from separate datasets show a decline in mortality rates. Cumulated data yielded longer, smoother trends with more narrow confidence intervals as opposed to the trends in the separate datasets. Unlike the trends from the separate datasets where the fluctuating rates might not be sufficient to infer a consistent pattern, trends from the cumulated dataset were more robust with a definite pattern that implies a decrease in mortality rates over time.

As shown in Figure 4, trends calculated from the separate datasets might contradict each other. For example, the NNMR and IMR calculated with the DHS 2015 survey data are subject to an upward sloping trend between 2000 and 2004. In contrast, calculating the NNMR and IMR with DHS 2014 survey data shows a decreasing trend between 2000 and 2004. Such inconsistencies might not suggest any significant differences across surveys, especially with the wide confidence intervals around the trends. This suggests the importance of cumulating survey data to produce reliable trends and to be able to correctly evaluate policies that are designed to reduce mortality rates.

Similar patterns are obvious in Figures A.1 and A.2 in the appendix where birth history data from Malawi DHS surveys (DHS surveys from 2000, 2004, 2010, and 2015) are cumulated with birth history data from Malawi MICS surveys (MICS surveys from 2006 and 2013).

Figure 4 Trends of childhood mortality rates based on a 5-year reference period in Senegal based on separate surveys and cumulated data



4.2 Childhood Mortality Rates: Subnational Level

Figure 5 reports trends for regional NNMR, IMR and U5MR of the following three regions in Senegal: Kaolack, Kedougou, and Sedhiou. Rates are again calculated based on separate survey datasets and a cumulated dataset. The separate survey rates, in the upper panel of the figure, are the rates calculated based on a reference period of 10 years before the survey and released in the DHS survey reports of DHS 2010-11, 2012-13, 2014, 2015, 2016, and 2017. The trends from the cumulated dataset are based on annual rates calculated based on two reference periods, 10 and 5 calendar years. Similar to the national trends, subnational trends from the cumulated dataset are longer and suffer from less fluctuation than trends from the separate datasets released by the DHS. The cumulation also allowed for rates based on a shorter reference period of 5 years as opposed to the standard 10-year period for subnational rates. The cumulated subnational rates are accompanied by standard errors that are less than or comparable to their counterparts from separate surveys, in case of the 10-year and 5-year rates, respectively.

Figure 6 reports trends of annual childhood mortality rates in the Kolda Region of Senegal, which were calculated on a reference period of 5 calendar years. The trends are for rates calculated from separate survey datasets and a cumulated survey dataset. As shown in Figure 6, cumulated data yields longer, smoother trends with more narrow confidence intervals. With the national rates, it is worth noting that some trends calculated with individual surveys contradict each other. For example, the NNMR calculated with DHS 2016 survey data was subject to an upward sloping trend between 2005 and 2009. In contrast, calculating the NNMR with DHS 2017 survey data showed a decreasing trend between 2005 and 2009.

Similar patterns are obvious in Figures A.3 and A.4 where birth history data from Malawi DHS surveys (DHS surveys from 2000, 2004, 2010, and 2015) are cumulated with birth history data from Malawi MICS surveys (MICS surveys from 2006 and 2013).

Figure 5 Trends of childhood mortality rates for selected regions in Senegal based on separate surveys and cumulated data

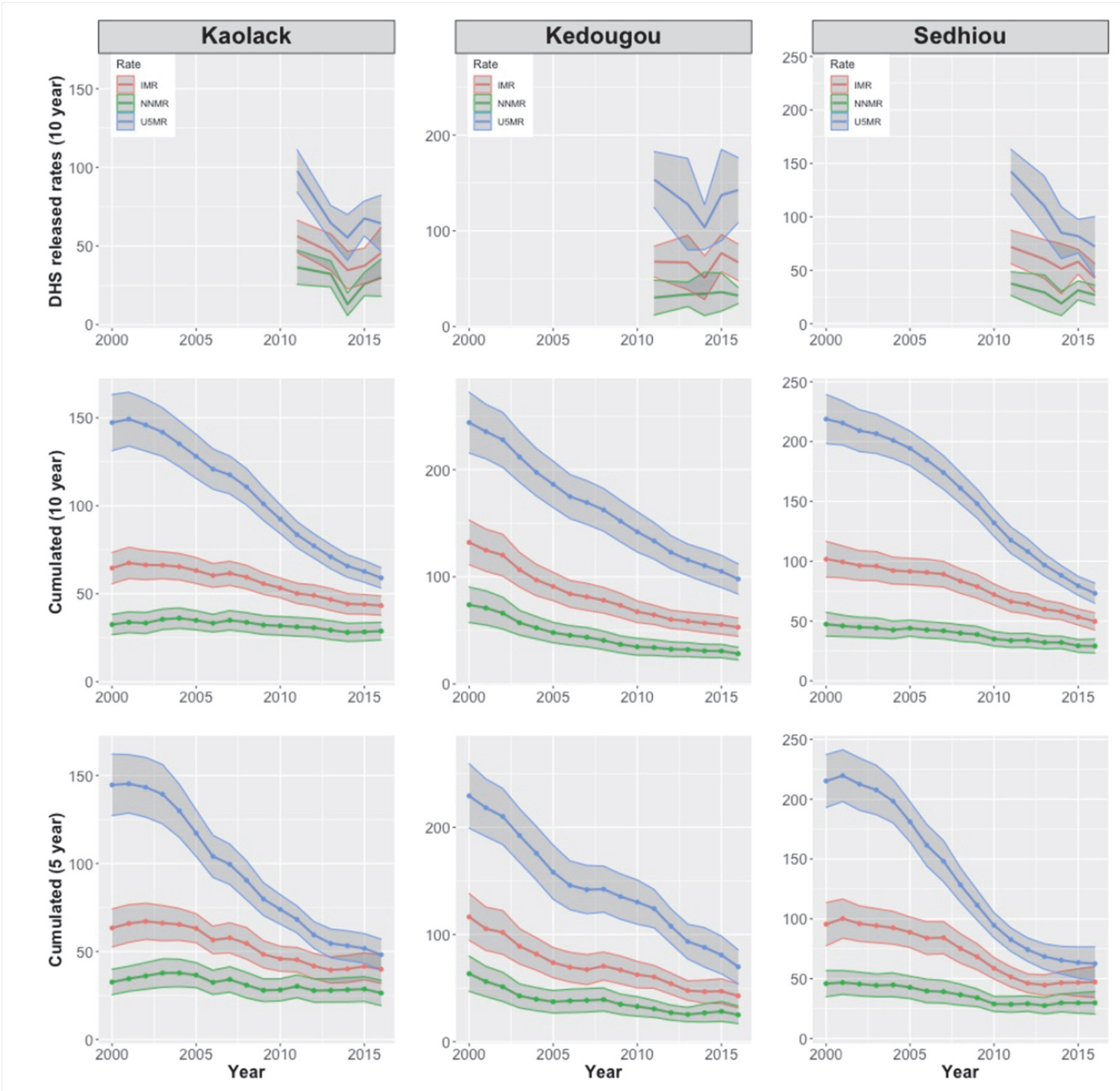
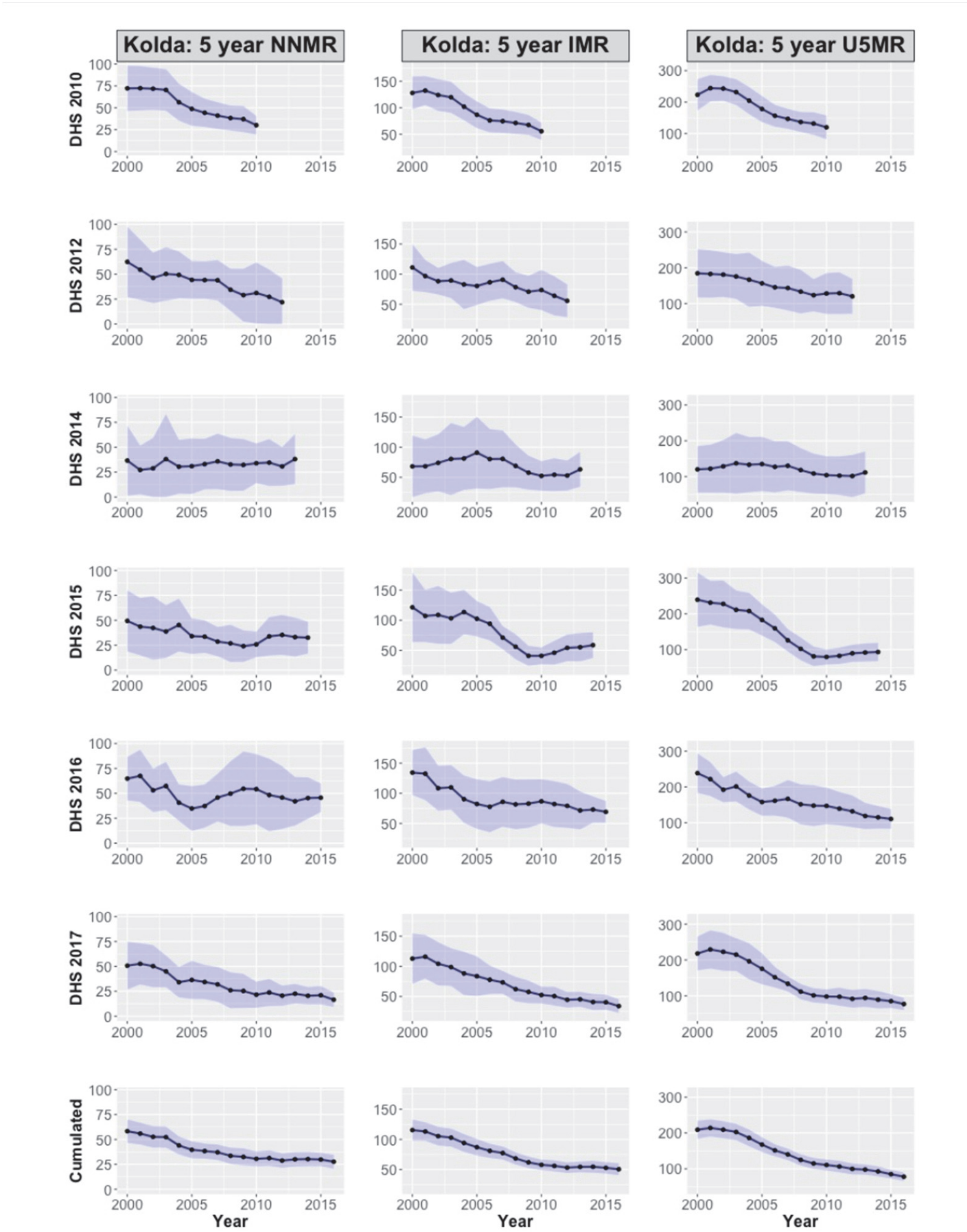


Figure 6 **Trends of childhood mortality rates based on a 5-year reference period in Kolda, Senegal based on separate surveys and cumulated data**



5 DISCUSSION

In this paper, we outlined a framework to guide the process of accumulating birth histories across different DHS/MICS surveys. This work aims to overcome the limited literature on a framework that guides the process for accumulating birth histories across several national DHS/MICS surveys, and utilizes the overlap in birth histories across such surveys. We focus on three rates that are essential for monitoring progress in meeting the UN SDGs: the NNMR, IMR, and U5MR. Based on DHS survey data for Senegal and DHS and MICS survey data for Malawi, we demonstrate that national and subnational trends based on mortality rates calculated with single survey data may suffer from fluctuations and might contradict each other when calculated from separate surveys.

We showed that using the proposed framework for accumulating birth histories across different surveys decreases the confidence intervals around the estimated mortality rates. Estimated standard errors of cumulated rates are 50-60% less than their counterparts from separate surveys. In addition, accumulating birth histories smoothes out fluctuations in time series for national and subnational mortality rates and establishes more stable, reliable time trends. It is important to acknowledge that the reliability of the cumulated time trend varies between national and subnational rates and with the reference period. Cumulated time trends for national rates based on a 5-year reference period and subnational rates based on a 10-year reference period appear to be very reliable when compared with cumulated rates based on a 1-year reference period or a 5-year reference period for national or subnational levels, respectively. Therefore, it is important to consider estimates based on shorter-than-standard reference periods with caution. Nevertheless, even with such limitations, the results stress the importance of using cumulating survey data to determine reliable trends and to correctly evaluate policies that are designed to reduce mortality rates.

REFERENCES

Abel, G. J., B. Barakat, K. C. Samir, and W. Lutz. 2016. “Meeting the Sustainable Development Goals Leads to Lower World Population Growth.” Proceedings of the National Academy of Sciences of the USA. 113(50): 14294–14299. https://doi.org/10.1073/pnas.1611386113.

Agence Nationale de la Statistique et de la Démographie (ANSD) [Sénégal], et ICF International. 2015. *Sénégal: Enquête Démographique et de Santé Continue (EDS-Continue 2012-14), Rapport Régional.* Rockville, MD, USA: ANSD et ICF International. http://www.ansd.sn/ressources/publications/Rapport%20EDS%20Continue%202015.pdf.

Chu, A., J. M. Brick, and G. Kalton. 1999. “Weights for Combining Surveys Across Time or Space.” *Bulletin of the International Statistical Institute*: *52nd Session, Contributed Papers, Book 2*: 103-4. https://www.stat.fi/isi99/proceedings/arkisto/varasto/kalt0185.pdf.

Croft, T. N., A. M. J. Marshall, C. K. Allen, et al. 2018. *Guide to DHS Statistics.* Rockville, MD, USA: ICF. https://www.dhsprogram.com/publications/publication-dhsg1-dhs-questionnaires-and-manuals.cfm.

Elkasabi M. 2020. DHS.rates: Calculates Demographic Indicators. R package version 0.7.1. https://cran.r-project.org/web/packages/DHS.rates/index.html.

Elkasabi M. 2019. “Calculating Fertility and Childhood Mortality Rates from Survey Data Using The DHS.Rates R Package.” *PLoS One* 14 (5): e0216403. https://doi.org/10.1371/journal.pone.0216403.

Elkasabi, M. A., S. G. Heeringa, and J. M. Lepkowski. 2015. “Joint Calibration Estimator for Dual Frame Surveys.” *Statistics in Transition New Series:* 16 (1): 7–36. https://www.academia.edu/28561958/Joint_Calibration_Estimator_for_Dual_Frame_Surveys.

Elkasabi, M. A. 2015. “Weighting Procedures for Dual Frame Telephone Surveys: A Case Study in Egypt.” *Survey Insights: Methods from the Field. Weighting: Practical Issues and ‘How to’ Approach.* 1-11. https://surveyinsights.org/wp-content/uploads/2015/02/Weighting-Procedures-for-Dual-Frame-Telephone-Surveys.pdf.

Hill, K. 2013. “Direct Estimation of Child Mortality from Birth Histories.” In *Tools for Demographic Estimation*, edited by T. A. Moultrie, R. E. Dorrington, A. G. Hill, K. Hill, I. M. Timæus, and B. Zaba. Paris: International Union for the Scientific Study of Population. http://demographicestimation.iussp.org/sites/demographicestimation.iussp.org/files/TDE_2013_2ndImpression.pdf.

Kalton, G. 2009. “Methods for Oversampling Rare Subpopulations in Social Surveys.” *Survey Methodology* 35 (2): 125-141. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.233.5247&rep=rep1&type=pdf.

Kish L. 1999. Cumulating/combining population surveys. *Survey Methodology* (Statistics Canada, Catalogue 12-001). 25 (2): 129-38. https://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/censocontinuo/Cumulating_Combining%20population%20surveys.pdf.

Korn, E. L., and B. I. Graubard. 1999. *Analysis of Health Surveys*. New York, NY, USA: Wiley. https://doi.org/10.1002/sim.773.

Lohr, S. 2011. "Alternative Survey Sample Designs: Sampling with Multiple Overlapping Frames." *Survey Methodology* 37 (2): 197-213. https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2011002/article/11608-eng.pdf?st=Eb3YFUY5.

MacQuarrie, K. L. D., W. Winfrey, J. Meijer-Irons, and A. Morse. 2018. *Consistency of Reporting of Terminated Pregnancies in DHS Calendars*. DHS Methodological Reports No. 25. Rockville, MD, USA: ICF. https://dhsprogram.com/publications/publication-mr25-methodological-reports.cfm.

National Statistical Office (NSO) and ICF. 2017. *Malawi Demographic and Health Survey 2015-16.* Zomba, Malawi, and Rockville, MD, USA: NSO and ICF. https://dhsprogram.com/pubs/pdf/FR319/FR319.pdf.

National Statistical Office (NSO) and ICF Macro. 2011. *Malawi Demographic and Health Survey 2010.* Zomba, Malawi, and Calverton, MD, USA: NSO and ICF Macro. https://dhsprogram.com/pubs/pdf/fr247/fr247.pdf.

National Statistical Office (NSO). 2015. *Malawi MDG Endline Survey 2014.* Zomba, Malawi: National Statistical Office. http://www.nsomalawi.mw/images/stories/data_on_line/demography/MDG%20Endline/MES%202014%20Report.pdf.

Roberts, G. and D. Binder. 2009. "Analyses Based on Combining Similar Information from Multiple Surveys." Survey Research Methods Section of the Joint Statistical Meetings (JSM), Washington, DC, USA, 2138-2147. https://www.semanticscholar.org/paper/Analyses-Based-on-Combining-Similar-Information-Roberts-Binder/70ef9ec615640c12f029fe0d026b0166ca88e2a7?p2df.

Pullum, T. and S. Assaf. 2016. *Long Term Trajectories of Fertility and Contraceptive Use*. DHS Analytical Studies No. 58. Rockville, MD, USA: ICF International. https://dhsprogram.com/pubs/pdf/AS58/AS58.pdf.

Thomas, S., and B. Wannell. 2009. "Combining Cycles of the Canadian Community Health Survey." *Health Reports* 20 (1): 53–58. https://www150.statcan.gc.ca/n1/pub/82-003-x/2009001/article/10795-eng.pdf.

United Nations, Department of Economic and Social Affairs, Population Division. 2011. *Mortality Estimates from Major Sample Surveys: Towards the Design of a Database for the Monitoring of Mortality Levels and Trends*. The Technical Paper series: Technical Paper No. 2011/2. https://www.un.org/en/development/desa/population/publications/pdf/technical/TP2011-2_MortEstMajorSampSurv.pdf.

United Nations Inter-agency Group for Child Mortality Estimation (UNIGME). 2018. *Levels & Trends in Child Mortality: Report 2018.* New York, NY, USA: United Nations Children's Fund. https://www.un.org/en/development/desa/population/publications/mortality/child-mortality-report-2018.asp.

Zambia Statistics Agency, Ministry of Health (MOH) Zambia, and ICF. 2019. *Zambia Demographic and Health Survey 2018*. Lusaka, Zambia, and Rockville, MD, USA: Zambia Statistics Agency, MOH, and ICF. https://dhsprogram.com/publications/publication-fr361-dhs-final-reports.cfm.

APPENDIX

Figure A.1 Trends of childhood mortality rates in Malawi based on separate surveys and cumulated data

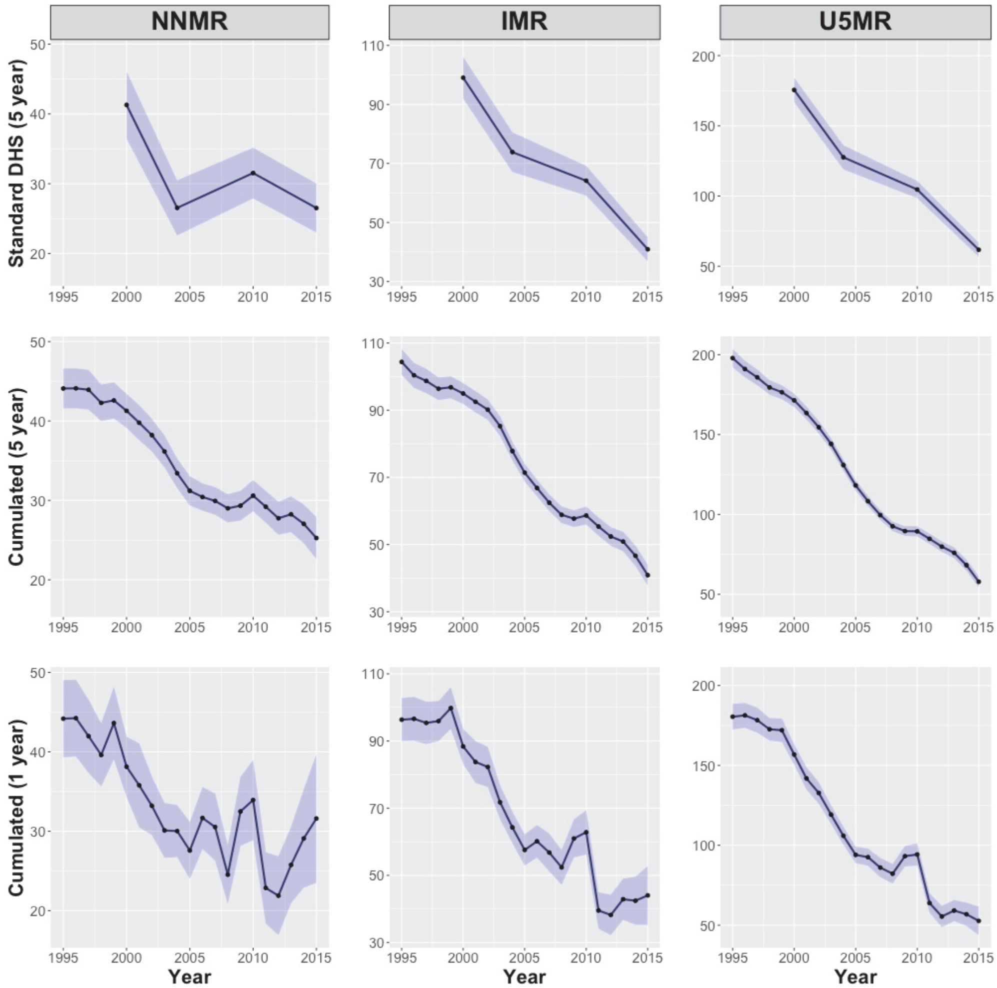


Figure A.2 Trends of childhood mortality rates based on a 5-year reference period in Malawi based on separate surveys and cumulated data

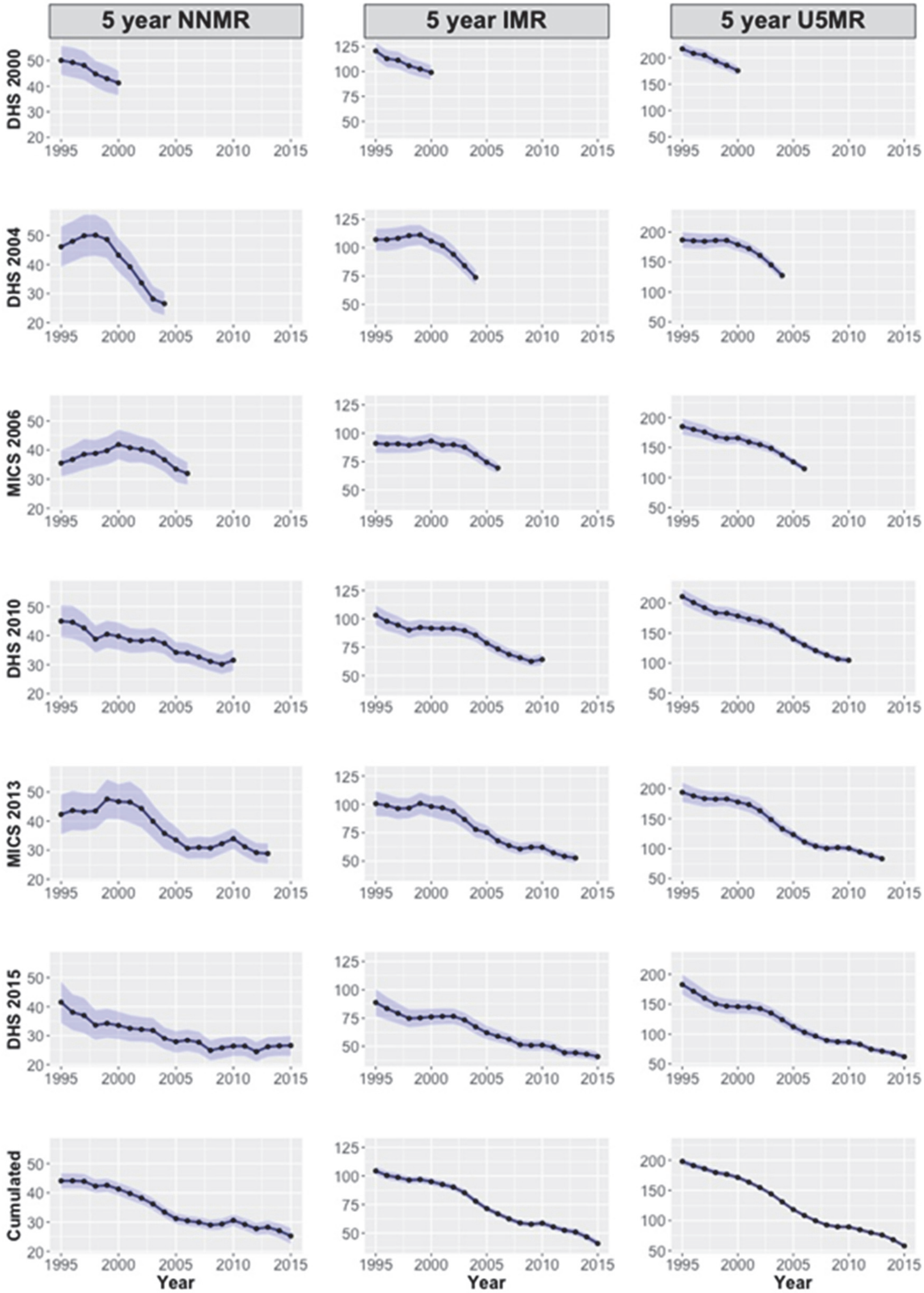


Figure A.3 Trends of childhood mortality rates for selected regions in Malawi based on separate surveys and cumulated data

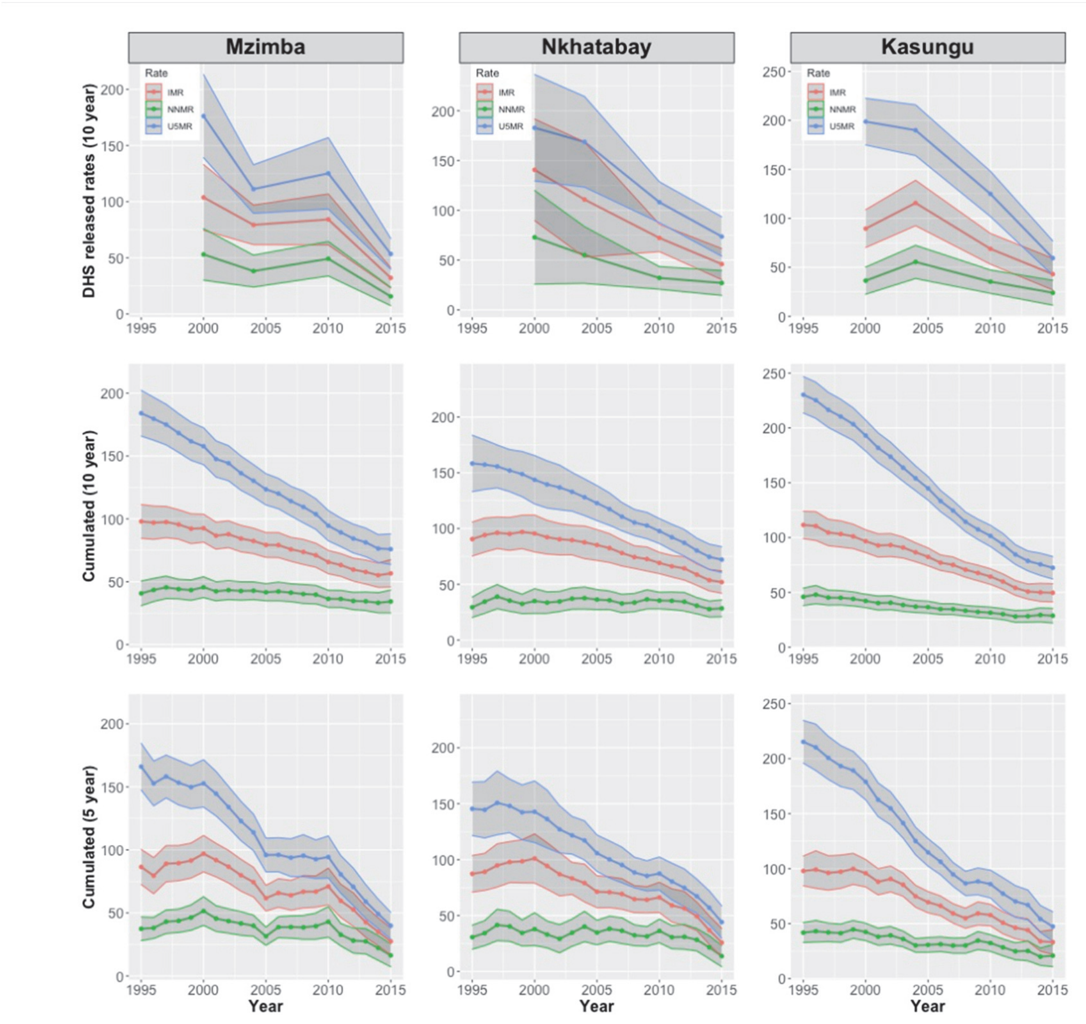


Figure A.4 Trends of childhood mortality rates based on a 5-year reference period in Blantyre, Malawi based on separate surveys and cumulated data