XXVII IUSSP International Population Conference

Busan, Korea, August 26-31, 2013

A New Approach to Indirect Estimation of Child Mortality: Application to Malawi

Agbessi Amouzou[1] and Kenneth Hill[1]

[1] Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe street, Baltimore, MD 21205

Agbessi Amouzou: aamouzou@jhsph.edu; Kenneth Hill: kenneth_hill_1@yahoo.com

Proposed session: **1st choice: 1705:** Indirect methods of mortality and fertility estimation: new techniques for new realities

Second choice: 306: New approaches to the collection and analysis of data on mortality and cause-of-death

Abstract:

Standard techniques of indirect estimation of child mortality use data from summary birth histories consisting of only two questions - number of children ever born alive and the number of children dead. However the estimation is based on several assumptions about fertility and mortality patterns, and rates computed for recent periods are biased. We propose and apply an innovative approach based on imputation of full birth histories onto summary birth histories. The resulting imputed full birth history is used to calculate child mortality rates using standard life table procedures. We apply the approach to data from the Malawi 2008 Population Census and the 2004 and 2010 Demographic and Health Survey datasets. Preliminary results are promising, with most of the imputed child mortality rates falling within the 95% confidence intervals of the rates directly computed from the 2010 DHS survey. In addition, choice of the existing full birth history data for the imputation did not appear to affect the resulting mortality rates computed from the imputed full birth history data.

Introduction

At the Millennium Summit in 2000, world leaders agreed on eight Millennium Development Goals, including a goal of reducing mortality among children less than five years of age by two-thirds between 1990 and 2015.[1] Monitoring this objective requires frequent estimates of under-five mortality levels at national level with sufficient precision to allow the detection of statistically significant trends. Although most low- and middle-income countries conduct surveys such as the USAID-supported Demographic and Health Survey (DHS)[2] every five years or so, these data are neither frequent nor recent enough to generate mortality estimates that can used to monitor short-run changes in under-five mortality needed to make improvements in child survival programs.[3] Given the slow pace of declines in under-five mortality, especially in poor countries, Governments and their development partners are advocating for "real-time" (annual) monitoring of child mortality as a strategy for tracking country progress. This strategy can only be implemented if good quality data are available every year for computation of annual under-five mortality rates. To date, low- and middle-income countries lacking complete civil registration systems have estimated childhood mortality directly by collecting full birth histories -- that is, dates of birth and for children that have died with ages at death for all live births -- from women aged 15-49 through household surveys, typically implemented through the DHS or UNICEF-supported Multiple Indicator Cluster Survey (MICS) programs. Collection of full birth histories in low- and middle-income countries requires lengthy and careful training of fieldworkers and close supervision in the field to ensure the high level of data quality needed for accurate mortality estimation. The burden of activities to be carried out during those surveys usually requires limiting the number of households sampled to fewer than would be required to generate annual mortality estimates with acceptable precision. In addition, because the surveys usually include many other modules, they are expensive to carry out every year. As an alternative to the full birth history, childhood mortality can be estimated indirectly using summary birth history data from women age 15-49 and applying the traditional methodology pioneered by William Brass.[4,5] A summary birth history consists of only two pieces of information collected from women aged 15-49: the number of children ever born and the number of children who have died. Additional information is obtained as an indicator of the average exposure of children to risk of death in the form of the age of the woman, the time since her first birth or the time since her first marriage. Estimates of child mortality are then derived from proportions of children dead, making assumptions about age patterns of fertility and child mortality.

Collecting a summary birth history is easier than collecting a full birth history, and requires less intensive training and supervision of fieldworkers. Summary birth histories can be included in national surveys designed for other purposes or in population census questionnaires, permitting frequent indirect estimation of child mortality at modest cost. Inclusion of summary birth histories in censuses would be particularly useful because it would support the estimation of mortality for small geographic areas. However, traditional indirect estimation methodologies do not provide reliable recent estimates of mortality because of bias, and they are heavily dependent on the choice of a particular mortality estimation model. These estimates are usually based on proportions dead of children ever born calculated for five-year age groups of mothers. The estimates do not refer to clearly-defined time periods, and the most recent mortality estimate obtained is based on information from women aged 15-19. Estimates based on reports from women in this age group are known to be positively biased due to selection bias; these young women are disproportionately likely to report first births and to have low socioeconomic status. To reduce this selection bias, other exposure variables such time since first birth or time since first union or marriage are used instead of the age of the mother. However, these variables are commonly more affected by data quality issues than age, and age-based mortality estimation is the most widely used.[6]

Due to these weaknesses in the indirect methods, countries have preferred the direct estimation method at the expense of timeliness and cost. Accurate real-time mortality monitoring requires innovative approaches. One such approach is to take advantage of the simplicity of collecting summary birth histories, thus allowing larger sample sizes and greater precision, while improving the analysis to produce estimates for clearly defined time periods.

We propose such an innovative approach to childhood mortality estimation and illustrate it with an application to Malawi, using summary birth histories and borrowing full birth history information from existing DHS surveys. Our hypothesis is that we can impute full birth history data from an existing survey onto summary birth histories to produce estimates of under-five mortality that fall within the confidence bounds of the estimates produced by the full birth histories.

Basis of birth history imputation

Although summary birth histories are often used for indirect mortality calculation, they cannot provide information about child mortality for a specific (short) recent time period. This is because the children whose births are reported by women are spread over a number of previous years, and are exposed to potentially time-varying mortality risks. The only exception would be children born to women who report their first birth as occurring in the 12 months before the interview, but this sample will be entirely first births and thus not representative of all children.

The birth history imputation approach we propose starts with a summary birth history, and borrows full birth history information from a pre-existing dataset such as a DHS. Individual women in the summary birth history dataset are matched to comparable women in the full birth history dataset based on the number of children ever born alive (CEB), the number of children who died among those born alive (CD) and age of the mother (or time since first birth if available in both datasets). Typically, the approach matches each woman who responded to the two summary birth history questions (woman “A”) to a woman with similar CEB, CD and age group (five-year age group) for whom a full birth history was collected in another survey – for example, a DHS survey (woman “B”). The full birth history details of woman “B” are then assigned to woman “A” to create a dataset with full birth histories which is then used to calculate under-five mortality. To reduce noise, each summary birth history is matched to multiple randomly-selected full histories with the same CEB/CD/Age characteristics. A relatively small proportion of women will not have an exact match on all three variables from the two datasets, in which case the match is made on CEB and CD only. Because the summary birth history data reflect any changes in overall mortality or fertility since the full birth history was collected (since the distribution of women by children ever born and children dead would change) the mortality estimates derived from this method should at least partially capture recent mortality and fertility trends. The key assumption of the approach is that the pattern of full birth histories for women with given CEB, CD and age does not vary over time.

The imputed full birth history derived in this way from a summary birth history survey is used to calculate childhood mortality using life table procedures identical to those used for a true full birth history dataset. The approach has several potential advantages compared to the traditional indirect method based on summary birth histories. First, it relies only on the assumption of no systematic temporal variation in the distribution of births and deaths given CEB, CD and age of the woman. Unlike the traditional indirect estimation methods, it does not rely on models of age patterns of fertility or mortality, and because it takes into account the actual distribution of women by age and CEB, is less affected by fertility trends. Second, because it uses information from all women 15-49 years, recent mortality estimates are not affected by the selection bias for young mothers observed in traditional indirect estimation. A final and most important advantage is its applicability to large datasets such as population census data, thus allowing estimation of recent or annual child mortality trends for small geographic areas. In addition, summary birth histories can be incorporated in national surveys designed for other purposes (e.g., SMART surveys, Welfare Monitoring Surveys, Integrated Household Surveys) and the data can be used to estimate annual mortality using the imputation approach.

Datasets used in Malawi application

We applied the birth history imputation approach using data from the 2008 Malawi Population and Housing Census (PHC) and the 2004 and 2010 Demographic and Health Surveys. The 2008 Malawi PHC included a summary birth history module that collected information on number of children ever born and number of children surviving from women aged 12 and over. We used a 10% sample of the PHC data, corresponding to 309,851 women 15-49, and took another 10% random sample of these women. Thus the total sample of women age 15-49 from the PHC data was 30,985. The 2004 and 2010 DHSs collected full birth histories for women aged 15-49. The 2004 survey collected data from 13,664 households and 11,698 individual women aged 15-49, whereas the 2010 survey collected data from 24,825 households and 23,020 individual women.[7,8] Imputed full birth histories for the 2008 PHC data were generated using both the 2004 and 2010 DHSs.

Analytical method

In each dataset, we categorized mothers by CEB, CD and five-year age group. Women in the PHC data set were randomly matched to women from one or other of the DHS datasets based on categories defined by CEB, CD and age. Because the PHC dataset had more mothers than the DHS dataset, we expanded the DHS data to obtain the same number of records within each category between the two datasets before the merge was conducted. A woman from the DHS category defined by CEB, CD and age group may therefore be matched to multiple women of the same category in the PHC data set. We repeated the imputation process ten times to generate imputed full birth history data, and repeated the process using first the 2004 DHS and second the 2010 DHS. Dates of birth in the full birth histories were adjusted to reflect the difference in interview date between these surveys and the census. We then computed annual neonatal, infant, and under-five mortality rates for each of the ten calendar years preceding the PHC data collection (1998-2007). We computed the same direct mortality rates using the

DHS 2010 datasets and compared the results with the imputed mortality rates. We estimated the 95% confidence intervals around the direct mortality rates computed from the 2010 DHS using the jackknife resampling method.

Imputation results

Neonatal mortality

Figure 1a presents the imputed neonatal mortality rates using summary birth history data from the 2008 Malawi PHC and borrowing full birth history data from the 2004 and 2010 DHSs, as well as the neonatal rates calculated directly from the 2010 DHS. Overall, the imputed neonatal mortality rates are fairly close to the direct rates, especially for the period from 1999 to 2002. Although the imputed neonatal rates are the same in 2003 for the two datasets, they slightly under-estimate the rates compared to the 2003 direct rate. There appears to be more discrepancies in recent periods. However, most imputed neonatal mortality rates fall within the 95% confidence intervals of the direct mortality rates computed from the DHS 2010. Figure 1b shows the ratio of each imputed rate to the corresponding directly calculated rate from the DHS 2010, showing two main results. First, the overall imputed neonatal rates are within 20% of the direct neonatal rates, except for the year 2006. Second, the imputed rates are very similar whether using the full birth histories from the 2004 DHS or the 2010 DHSs, especially between 1998 and 2003, indicating that the choice of DHS data set does not make much difference to the results, at least in this instance.

Infant mortality rates

Figure 2a presents the imputed infant mortality rates using the full birth histories from the 2004 and 2010 DHSs with the 2008 census summary histories, and the direct estimates from the 2010 DHS. The imputed infant mortality rates are very close to one another and to the direct estimates from the 2010 DHS, and falling mostly within the 95% confidence interval of the direct mortality rates. The imputed rates using the 2004 DHS are particularly close to the direst estimate from the 2010 DHS and reflect almost perfectly the trends observed in the 2010 DHS. The results are confirmed in Figure 2b, showing the ratios of the imputed rates to the direct rates. Almost all imputed rates using the 2004 DHS fall within 10% of the direct infant mortality rates from the 2010 DHS, and the imputed rates using the 2010 DHS fall within 20% of the direct rates.

Under-five mortality rates

Figures 3a and 3b show the imputed under-five mortality rates and the ratio of the imputed rates to the direct 2010 DHS rates. Similar to the infant mortality rates, the imputed rates are very close to one another and to the direct estimates. Except for the year 2003, all the imputed mortality rates fall within the 95% confidence intervals of the direct under-five mortality rates. The imputed rate observed in 2003 appears to depart slightly from the general trend, but with this exception all imputed under-five mortality rates fall within 20% of the direct estimates.

Figure 1a: Imputed neonatal mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 and 2010 DHS, and direct neonatal mortality estimates from the 2010 DHS with 95% confidence intervals, 1998-2007

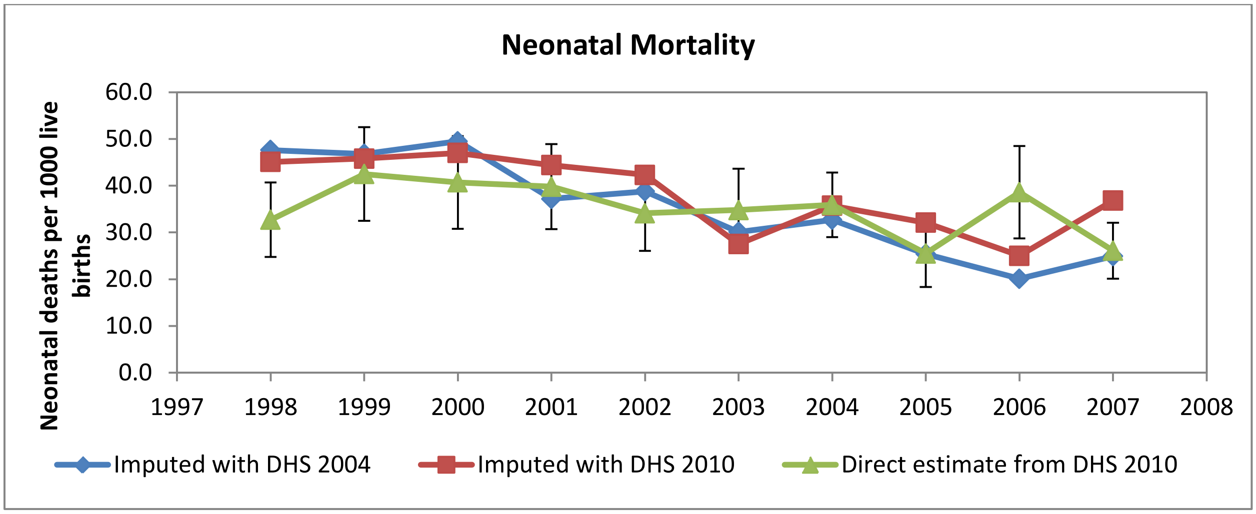


Figure 1b: Ratios of imputed neonatal mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 and 2010 DHS, to direct neonatal mortality estimates from the 2010 DHS

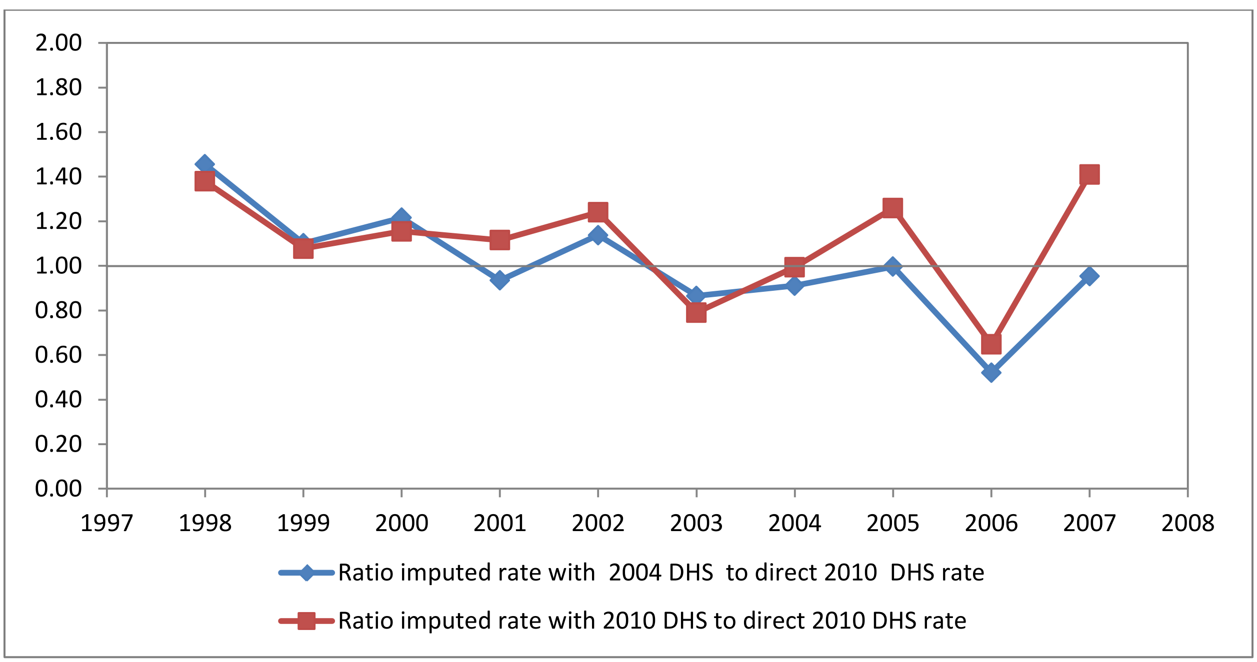


Figure 2a: Imputed infant mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 ad 2010 DHS, and direct infant mortality estimates from the 2010 DHS with 95% confidence intervals, 1998-2007

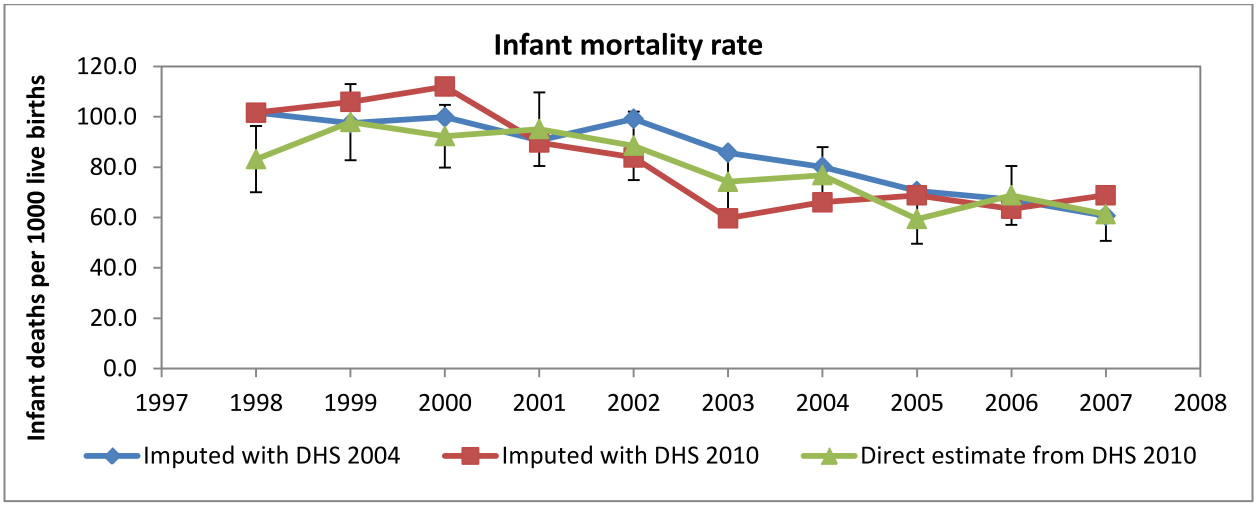


Figure 2b: Ratio of imputed infant mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 ad 2010 DHS, to direct infant mortality estimates from the 2010 DHS

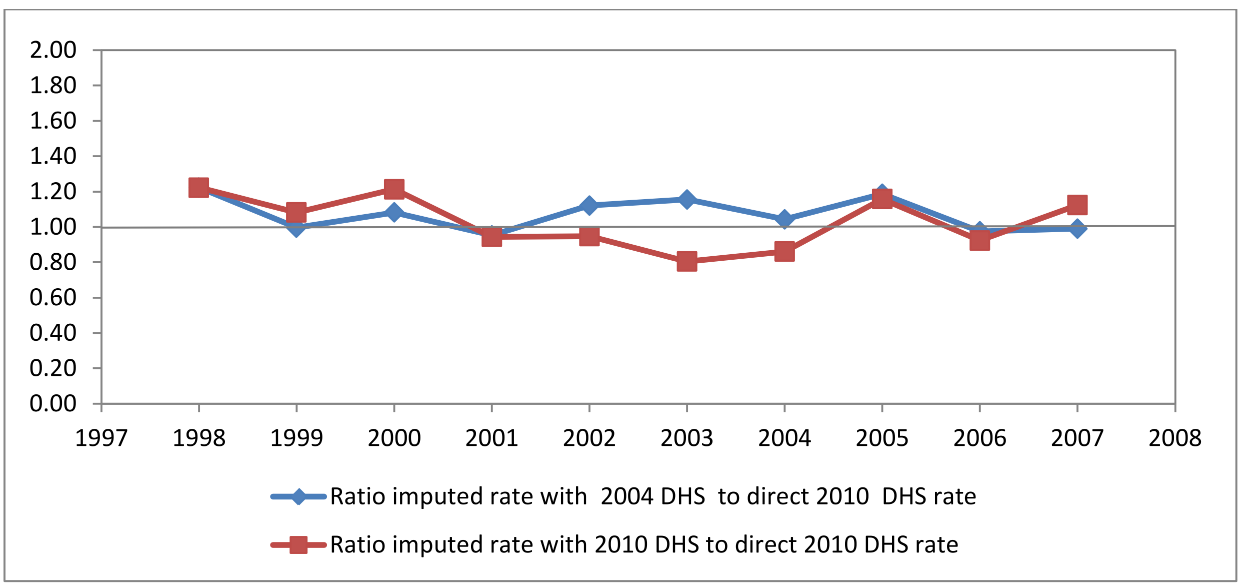


Figure 3a: Imputed under-five mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 and 2010 DHS, and direct under-five mortality estimates from the 2010 DHS with 95% confidence intervals, 1998-2007

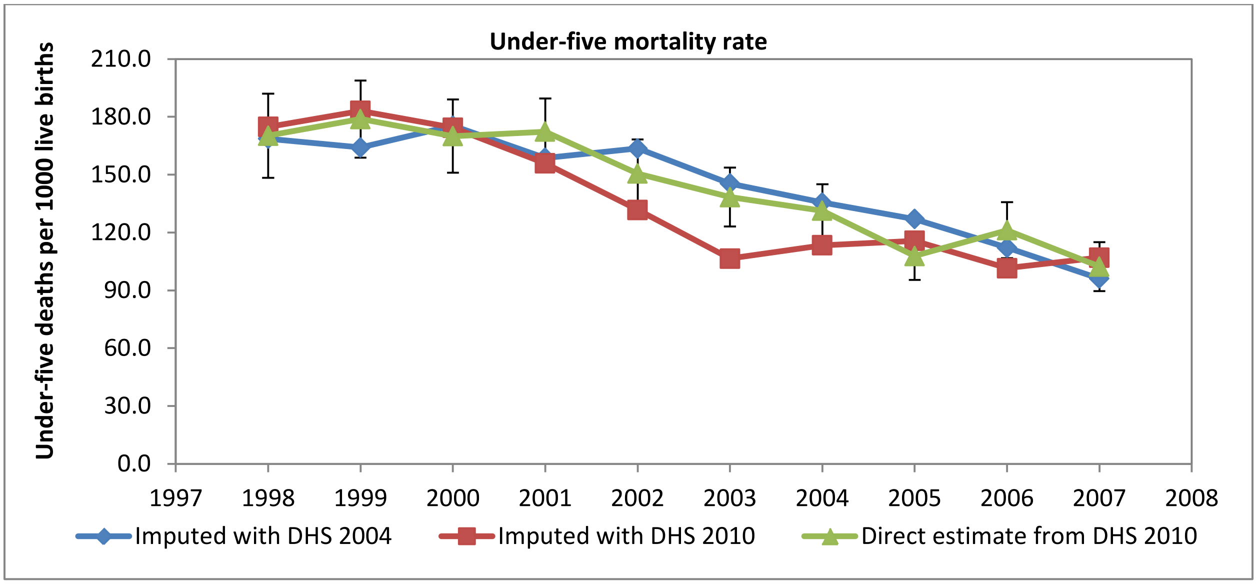
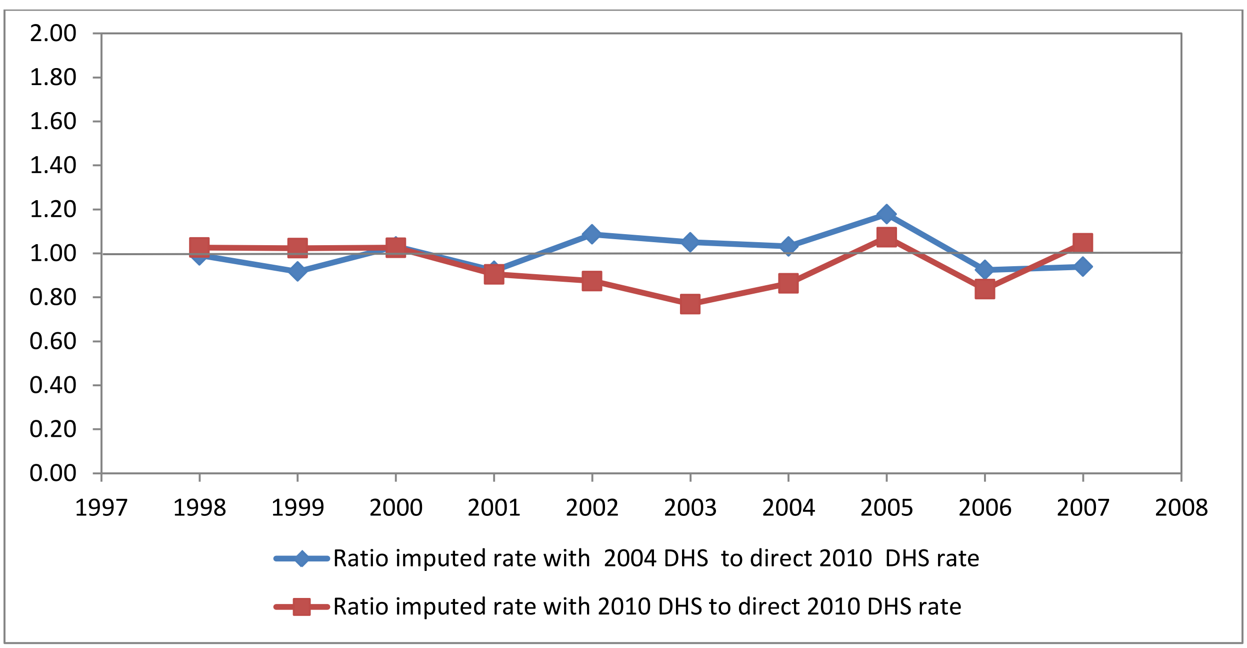


Figure 3b: ratios of imputed under-five mortality rates using summary birth history data from the Malawi 2008 PHC and borrowing full birth history data from the 2004 and 2010 DHS, to direct under-five mortality estimates from the 2010 DHS



Discussion

The preliminary results of this application of the imputation approach are promising. Annual neonatal, infant and under-five mortality rates based on the imputed data are very close to the corresponding rates directly computed from the 2010 DHS, with almost all rates falling within 95% confidence intervals of the direct rates. Despite fluctuations due to the smaller sample sizes available for calculation of annual rates, the approach appears to capture both trends and levels of mortality reasonably well. The fact that the method performed equally well using either the 2004 or the 2010 DHS as the source of full birth history data is also promising, suggesting that the method is sufficiently robust to be used with different types of surveys over variable time periods. Further applications are needed to confirm the favorable characteristics of the method. Once fully validated, the imputation approach offers important benefits for countries seeking "real-time" information on trends in under-five mortality. Summary birth histories can be included in surveys that are conducted annually by national statistics offices, producing data that can be used to compute child mortality for recent years at very low cost. In countries where summary birth histories are included in population censuses, this method can be used to estimate mortality for small geographic areas such as districts or sub-districts and for short reference periods.

References

[1] United Nations. The Millennium Development Goals Report 2011. New York: United Nations.

[2] Available: www.measuredhs.com. Accessed 20 October 2012.

[3] Korenromp EL, Arnold F, Williams BG, Nahlen BL, Snow RW. Monitoring trends in under-5 mortality rates through national birth history surveys. *International Journal of Epidemiology* 2004;33:1293–1301.

[4] Brass W. Use of census or survey data for the estimation of vital rates. E/CN.14/CAS.4/57. *African Seminar on Vital Statistics*, Addis Ababa, 14-19 December, 1964.

[5] Brass W. Methods for estimating fertility and mortality from limited and defective data. Technical report, Carolina Population Center, Laboratories for Population Statistics. Chapel Hill, North Carolina, 1975.

[6] United Nations. Manual X. Indirect techniques for demographic estimation. *Department of International Economic and Social Affairs. Population Studies*, No 81. New York, 1983.

[7] National Statistical Office (NSO) and ICF Macro. 2011. Malawi Demographic and Health Survey 2010. *Zomba, Malawi, and Calverton*, Maryland, USA: NSO and ICF Macro

[8] National Statistical Office (NSO) [Malawi], and ORC Macro. 2005. Malawi Demographic and Health Survey 2004. *Calverton, Maryland*: NSO and ORC Macro.