Child mortality estimation in countries with high HIV/AIDS prevalence.

Reference year of estimates: 2008

By the Inter-agency Group for Child Mortality Estimation (IGME), 2009

Estimation methods of child mortality in countries with high prevalence of HIV/AIDS require adjustment to survey data to account for possible biases in data collection, and modifications to standard implementation of statistical regression methods to capture the extraordinary rapid changes in child mortality over the HIV/AIDS epidemic period. The general approach is to subtract estimated HIV/AIDS deaths from empirical data points of child mortality, then to use statistical regression modeling to fit a regression curve estimating the trend in non-HIV/AIDS child deaths, then finally to add back the estimated time series for HIV/AIDS deaths.

Estimates of child mortality for populations lacking accurate registration of births and deaths are almost entirely derived from reports of mothers about the survival of their children from household survey such as the Demographic and Health Survey (DHS). In populations severely affected by HIV/AIDS, HIV positive (HIV+) children will be more likely to die than other children, and will also be less likely to be reported since their mothers will have been more likely to die also. Child mortality estimates will thus be biased downwards. The magnitude of the bias will depend on the extent to which the elevated under-five mortality of HIV+ children is not reported because of the deaths of their mothers.

The Technical Advisory Group (TAG) of the Inter-agency Group for Child Mortality Estimation developed a method to adjust HIV/AIDS related mortality for each survey data observation during HIV/AIDS epidemic (1980-present), by adopting a set of simplified but reasonable assumptions about the distribution of births to HIV+ women, primarily relating to the duration of their infection, vertical transmission rates, and survival times of both mothers and children from the time of the birth (see Annex 1 for a more detailed description of this methodology).

After adjusting the survey data for HIV/AIDS related survivor bias in the epidemic period, UNAIDS-WHO estimates of HIV/AIDS under-five mortality are subtracted from all data points in the epidemic period. One of two standard approaches is then used to fit a regression curve to the data for non-HIV deaths. The regression curve is extrapolated to year 2008 if the most recent data point is older than 2008. Finally, the UNAIDS-WHO estimates of HIV/AIDS under-five mortality are added to estimates from the regression model for the epidemic period. Annex 2 describes the statistical regression modeling.

Infant mortality rate (IMR) estimates are derived from under-five mortality rate (U5MR) by selecting appropriate model life tables. For the epidemic period, a non-HIV curve of infant mortality is derived from the non-HIV under-five mortality curve and then the UNAIDS estimates of HIV/AIDS deaths for children under age 1 are added to generate the final infant mortality estimates.

The child mortality estimation methods described in this document were used for 17 countries with high prevalence of HIV/AIDS as defined by a prevalence rate exceeding 5% at any point in the epidemic period. The 17 countries are: Botswana, Cameroon, Central Africa Republic, Côte d'Ivoire, Gabon, Kenya, Lesotho, Malawi, Mozambique,

Namibia, Rwanda, South Africa, Swaziland, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe. The estimates of under-five mortality rate and infant mortality rate produced by IGME are listed in Table 1.

Table 1. Under-five mortality rate (U5MR) and infant mortality rate (IMR)

	U5MR						IMR					
Country	**1990**	**1995**	**2000**	**2005**	**2007**	**2008**	**1990**	**1995**	**2000**	**2005**	**2007**	**2008**
Botswana	50	62	81	36	32	31	39	43	55	29	26	26
Cameroon	149	149	147	138	133	131	92	93	91	86	83	82
Central African Republic	178	180	181	177	175	173	116	118	119	117	116	115
Côte d'Ivoire	150	147	138	122	116	114	104	102	96	86	82	81
Gabon	92	89	87	82	78	77	67	65	64	60	58	57
Kenya	105	122	128	127	127	128	68	77	81	80	80	81
Lesotho	101	103	109	97	85	79	80	80	83	74	67	63
Malawi	225	203	162	123	107	100	133	122	100	78	69	65
Mozambique	249	223	183	149	136	130	166	149	124	103	94	90
Namibia	72	71	77	58	43	42	49	48	51	41	32	31
Rwanda	174	225	186	136	119	112	106	132	112	85	76	72
South Africa	56	57	73	70	66	67	44	43	52	50	47	48
Swaziland	84	93	124	95	89	83	62	66	83	66	62	59
Uganda	186	174	158	143	138	135	114	107	98	89	86	85
United Republic of Tanzania	157	163	139	117	108	104	97	100	87	74	69	67
Zambia	172	180	169	158	151	148	105	110	104	97	94	92
Zimbabwe	79	91	102	99	97	96	51	55	62	63	62	62

Note that birth transference is also a serious issue in many DHS surveys conducted in the countries with high HIV prevalence. Birth transference basically refers to the date of birth being incorrectly reported or recorded as occurring earlier than is actually the case.[1] In DHS surveys, this birth transference tends to be more pronounced for deceased than for surviving children. When this occurs it results in the under-five mortality rate being under-estimated for the most recent period (0 – 4 years before the survey). The adjustment for birth transference was done before the adjustment of HIV/AIDS related deaths for data from the surveys with serious birth transference issues.

Summary of analysis steps for under-five mortality rate estimation in countries with high HIV/AIDS prevalence:

1. Adjust survey data for birth transference.
2. Adjust survey data for HIV/AIDS related survivor bias in epidemic period as described in Annex 1.
3. Exclude indirect data series from 1990 onwards if age group of mothers is less than 25 or more than 39 years.

[1] Birth transference is partly caused by the design of the questionnaire of a survey. DHS questionnaires include a lengthy series of questions which are asked to mothers concerning maternal and child health. This series of questions must be asked for all births for which the date of birth is subsequent to a specified date—usually set as January of the fifth or sixth calendar year proceeding the year of the survey. It appears that interviewers learn that they can reduce their workload by incorrectly recording some births that actually occurred after the cutoff date as occurring prior to that date. In DHS surveys, interviewers appear to be particularly anxious to avoid asking the health questions about deceased children.

4. Evaluate whether indirect data series from 1990 onwards for age group of mothers 25-39 years are to be included:
   A. Fit a regression curve, as outlined in Annex 2, with indirect data series with survey year 1990 onwards excluded.
   B. Exclude indirect data series with survey year 1990 onwards if two or more of the survey estimated U5MR for age groups 25-39 are greater than regression model estimates for the same years.
5. Fit regression curve to all remaining data points as described in Annex 2.

Annex 1
Adjusting child mortality estimates from survey data in high HIV populations

The problem

Estimates of child mortality for populations lacking accurate registration of births and deaths are almost entirely derived from reports of mothers about the survival of their children. The most common approach is the full birth history (used in all Demographic and Health Surveys and some Multiple Indicator Cluster Surveys), whereby each woman aged 15 to 49 at the time of the survey is asked for the date of birth and, if the child has died, age at death of each live-born child she has had. Child mortality is then estimated from the information on births and deaths for periods up to 25 years before the survey. To be valid estimates of population-level child mortality, the births must be representative of all births in the population for the entire period. It is unlikely that this condition is ever perfectly met: births to mothers who have died or migrated out of the population will not be reported, and for periods long before the survey births to older mothers will not be represented because the mothers will have been age 50 or over at the time of the survey. It is normally assumed that any bias introduced by lack of representativeness is small, but this will not be the case in populations severely affected by HIV where HIV-positive (HIV+) children will be more likely to die than other children, and will also be less likely to be reported since their mothers will have been more likely to die also. Child mortality estimates will thus be biased downwards.

The adjustment methodology

The magnitude of the bias will depend on the extent to which the elevated under-five mortality of HIV+ children is not reported because of the deaths of their mothers. Precise estimates of the bias would require a great deal of information about the HIV epidemic, such as the distribution of births to HIV+ women by the duration of their infection, vertical transmission rates, survival times of both mothers and children from the time of the birth, and much else besides. Such information is not typically available, so we have adopted a number of simplifications.

We start with the latest information of a national population and its HIV epidemic from UNAIDS. UNAIDS typically provides the annual number of births from 1970 onwards, and the HIV prevalence among pregnant women aged 15 to 49. For each year, the births are divided into three components: births to HIV-negative (HIV-) mothers, assumed to be all HIV-; HIV- births to HIV+ mothers; and HIV+ births to HIV+ mothers (no distinction is made between children infected at or before birth and those infected after birth). For births in each year, deaths under age 5 in the subsequent five years are calculated for each component. For the HIV- births, risks of dying are obtained from a model life table (Coale-Demeny "West" family) with a U5MR approximating a best estimate as to U5MR in the HIV- population. Although there is growing evidence that mortality risks increase for HIV- children if their mothers are HIV+, this factor was not taken into account in the model. HIV- children are assumed to have the same risk regardless of the HIV status of the mother. For HIV+ births, a mortality schedule derived from cohort studies, with a probability of dying by age 5 of 62.5% is used; it is assumed that antiretroviral treatment has had no effect on these risks for the relevant period.

We now have estimates of true births and under-five deaths for each calendar year. The next step is to estimate how many of these births and under-five deaths will go unreported at a

particular survey because of deaths of the mothers. We first assume that HIV- women have no mortality, so all their births and under-five deaths are reported. We then assume that births to HIV+ women occur on average to women after 4 years of infection (4 years was chosen because it allows for a reduction of fertility among women who are HIV+ of approximately 25%). A survival curve from first infection, again derived from cohort studies, with a median survival time of about 9.5 years, was used to create a survival curve from 4 years after infection, and the probabilities of surviving from a particular year to the year of a given survey (assumed to be at the end of a year) were used to calculate the proportion of the births and child deaths (whether HIV- or HIV+) of HIV+ mothers reported by the survey.

For each 5 year period before a survey, the true and the "reported" births and under-five deaths are summed, and the estimated bias is calculated as the ratio of the reported under-five deaths divided by the reported births to the corresponding value for the true numbers. Survey estimates of under-five mortality are then adjusted by dividing by the bias estimates for each time period.

An example will make the process clearer. Assume a survey was conducted at the end of 2005, and consider births that occurred in 1995. UNAIDS provides the number of such births, and the HIV prevalence in 1995 among pregnant women, from which we derive the number of births to HIV- and HIV+ women. Births to HIV+ women are then subdivided on the assumption that they were 65% HIV- and 35% HIV+. We estimate an appropriate level of U5MR for HIV- births (largely from estimates prior to the epidemic), let's say 150 per 1,000 live births, roughly speaking appropriate for Tanzania. Under-five deaths will occur in the period 1995 to 2000; all the births and under-five deaths to HIV- women will be reported, but very few of the births and deaths (regardless of the serostatus of the child) to HIV+ mothers will be reported (the 10-year survival probability from 4 years after infection is assumed to be only 14 percent). True births and under-five deaths and reported births and under-five deaths are summed for the periods 2001-2005, 1996-2000, etc., and for each period the ratio of reported mortality rate to true mortality rate is calculated. Adjusted survey estimates for each period are then obtained by dividing the observed values by the ratio just calculated.

Annex 2
Fitting a regression curve to U5MR data points

IGME evaluated two statistical regression models for each country of the world and applied the best fitting model to estimate child mortality trends over the last decades till 2008. The regression curve was extrapolated to year 2008 if the most recent data point was older.

The two regression models used by IGME are the Spline regression[2] and Loess regression[3]. Table 2 outlines which of the two regression models was used for countries with high prevalence of HIV/AIDS. To capture the extraordinary rapid changes in child mortality driven by HIV/AIDS over the epidemic period in these countries, UNAIDS-WHO estimates of HIV/AIDS deaths under 5 were used as inputs[4]. The regression models were fitted to estimated non-HIV/AIDS deaths using data points from which the HIV/AIDS deaths had been subtracted. Subsequently, the UNAIDS-WHO estimates of HIV/AIDS under-five mortality were added to estimates from the regression model.

Table 2. Statistical regression models used

Country	Regression model
Botswana	Spline
Cameroon	Loess
Central African Republic	Loess
Côte d'Ivoire	Loess
Gabon	Loess
Kenya	Spline
Lesotho	Loess
Malawi	Spline
Mozambique	Spline
Namibia	Loess
Rwanda	Spline
South Africa	Loess
Swaziland	Loess
Uganda	Loess
United Republic of Tanzania	Spline
Zambia	Loess
Zimbabwe	Loess

The steps in the implementation of the regression models for the countries with high prevalence of HIV/AIDS were:

1. Subtract UNAIDS-WHO estimates of HIV/AIDS under-five mortality from data points in the epidemic period. Outside the epidemic period child mortality from HIV/AIDS is minor and so step 1 results in a data set assumed to represent under-five mortality from all other causes than HIV/AIDS.
2. Fit regression model to all data points for under-five mortality from all other causes than HIV/AIDS.

[2] Inter-agency Group for Child Mortality Estimation. *Levels and Trends of Child Mortality in 2006* [Working Paper], New York, 2007 (www.childinfo.org/files/infant_child_mortality_2006.pdf).

[3] Murray CJ et al. *Can we achieve Millennium Development Gold 4? New analysis of country trends and forecasts of under-5 mortality to 2015*. Lancet 2007, 370(9592), 1040-54

[4] UNAIDS-WHO Working Group on global HIV/AIDS and STI surveillance. (http://www.unaids.org/en/KnowledgeCentre/HIVData/Epidemiology/epiworkinggrp.asp)

3. Extrapolate regression curve to 2008.
4. Add UNAIDS-WHO estimates of HIV/AIDS under-five mortality to regression estimates in the epidemic period.

Botswana was the exception to this implementation of regression models. To better capture changes in child mortality in Botswana over the last decades, a Spline regression curve was fitted to data points for all-cause under-five mortality. From the most recent data point till year 2008, under-five mortality from all other causes than HIV/AIDS was assumed constant to the regression estimated under-five mortality for the most recent data point. The regression curve was extrapolated till 2008 by adding UNAIDS-WHO estimates of HIV/AIDS under-five mortality.