THE INFANT MORTALITY FRACTION AND ASSESSMENT OF MORTALITY DATA IN REFUGEE CAMPS: A Comparative Study Using Data from HMD, DHS and UNHCR

Courtland Robinson, Yoonjoung Choi and Vladimir Canudas-Romo

A. INTRODUCTION

In most developed countries, death rates are monitored by civil registration systems that record vital statistics like birth, death, marriage and divorce. In many developing countries, however, vital statistics systems are of poor quality or lacking altogether. Each year, nearly 50 million newborn children are not registered and only a third of countries outside of Europe and North America have the capacity to obtain usable mortality statistics (Setel et al, 2007). Even if they were once functioning, vital statistics systems often are disrupted in situations of civil conflict or natural disaster. In some cases, it may be possible to establish registration systems—in refugee camps or internally displaced persons (IDP) settlements, for example—to record vital events and other health indicators. In other instances, surveillance systems may offer the best approach to tracking morbidity and mortality prospectively (Checchi and Roberts, 2008). In many crisis settings, research on mortality must rely on cross-sectional surveys to estimate death rates retrospectively (Prudhon and Spiegel, 2007).

While mortality is the final and most definitive health outcome in almost any setting, understanding—and estimating—mortality in a crisis is even more important. First, knowing how many people have died (or, more accurately, the rate at which people have died) may help to determine how urgent is the need for humanitarian action. Second, knowing something about who has died—in terms of age, sex, cause of death, and other characteristics—helps to determine who is at greatest risk in the surviving population. Third, mortality rates (particularly of children under five) have proven to be sensitive tracking indicators of overall population health and security and thus may be useful in monitoring changes in health over time.

While measuring mortality in crisis settings is vital, measurement itself is hampered by myriad complicating factors: the need for rapid data, limited resources (including human and financial), insecure environments, and volatile population movements. Mortality data gathered in these complex and constrained settings present challenges to those who must review and use them, as it is not often clear whether wide variations reflect true heterogeneity of crisis type, severity, and context or whether they may also be due to data problems and errors.

In order to assess the quality and plausibility of data gathered in resource-poor crisis settings, it may be useful to start by examining mortality data of reasonably good quality—even and particularly when that data include periods of crisis (war, flu pandemic, etc)—and then to compare that to data of possibly lower quality, including from developing countries experiencing crisis mortality. For purposes of this study, we have identified as our “gold standard,” data from the Human Mortality Database (HMD) including data from 37 European countries, spanning a period of four centuries. Our comparison data for contemporary developing countries come from 161 Demographic and Health Surveys (DHS) conducted in 61 countries between 1985 and 2007. It is suggested the surveys have potential data quality problems in mortality estimation, but they were conducted in populations/situations that were stable enough to conduct a large-scale national survey. Finally, we used refugee camp data collected by the United Nations High Commissioner for Refugees (UNHCR) in roughly 80 camps spanning 16 countries, primarily in Africa and Asia (UNHCR, 2009).

The primary purpose of the paper is to compare age-patterns of under-five mortality—specifically the infant mortality fraction, which is a ratio of infant mortality (1q0) to under-five mortality (5q0)—relative to the overall level of under-five mortality across populations measured in these three data sets. Three main specific aims are:

1. To explore the relationship between infant mortality and under-five mortality in non-crisis periods and in crisis periods (specifically, the flu pandemic of 1918 and World War II) using historical data from developed countries, using the HMD data.
2. To compare the relationship in contemporary developing countries, using data from the DHS, to that observed in the HMD data; and to assess differential relationships by region and data quality.
3. To compare the relationship observed in health information systems data obtained by the UNHCR to that observed in the HMD data; and to assess differential relationships by region, camp, and other characteristics (including age distribution, age of camp, migration patterns, data collection systems, and health indicators).

B. DATA

1. Human Mortality Database

The Human Mortality Database (HMD, 2009) contains detailed time series of mortality data and life tables for populations with virtually complete registration and census data. Currently there are 37 countries/areas in the database. HMD contains 719 period life tables covering (mostly) five-year time intervals spread across four centuries. All life tables in this collection were computed directly from observed deaths and population counts, without adjustment except at the oldest ages. All of these five-year period life tables from the HMD are utilized in the present analysis, except for the data from Iceland, Luxemburg and Taiwan. Results from Iceland and Luxemburg showed significant variability due to their small population counts, and there are problems with the initial series of Taiwanese data. To study changes in the infant fraction during the flu pandemic of 1918 and during the war years we have also employed annual life tables from HMD. Probabilities of dying before age one and before age five are extracted from each of these HMD period life tables.

2. Demographic and Health Surveys

We selected 161 Demographic and Health Surveys (DHS) conducted in 61 countries between 1985 and 2007, which will be analyzed for the study. The DHS, which are nationally representative household surveys, have collected data on the population, health, and nutrition of women and children. The surveys collect full birth histories from 15 to 49 year old women in sampled households, providing data for direct estimation of childhood mortality. The respondent is asked to report all children ever born and full birth history for each live-born child. The birth history data include information about the date of birth, survival status, and age at death, if dead, of each child. Age at death is reported in days for deaths in the first 28 days of life, in months for deaths between 1 and 23 months, and in years for deaths at ages 2 and over. These data provide a basis for direct estimation of child mortality, specific by age and time period.

3. UNHCR Refugee Camp Health Information System Data

The UNHCR Health Information System (HIS) is a standardized tool to monitor and evaluate refugee public health programs and is currently designed to monitor stable, camp-based populations. As of early 2009, a total of 16 countries in Africa and Asia were reporting into the

HIS. The total population under surveillance was nearly 1.4 million across 80 camps (UNHCR, 2008). For this analysis, we used monthly data for 2008 (the most recent complete year) and calculated age-specific mortality rates using the average annual camp population and the reported total number of deaths in 2008 per camp and by age group in year (0 and 1-4).

C. METHODS

1. Measurement

a. Age patterns of childhood mortality: infant mortality fraction

An infant death and an under-five death refer to any death between 0-11 and 0-59 completed months of age, respectively. The probability of dying before age one year of age for a birth cohort, denoted as ${}_1q_0$, is calculated as the ratio of deaths under age one divided by the number of births in that cohort. Similarly, the probability of dying before age five years of age, denoted as ${}_5q_0$, is calculated as the ratio of under-five deaths divided by the number of births in that cohort. In a period life table, such as in HMD period life tables, ${}_1q_0$ is equal to the fraction of infant deaths of the total deaths, and ${}_5q_0$ is equal to the fraction of under-five deaths of the total deaths. These measures are readily available in the HMD life tables (HMD, 2009). Using birth histories from DHS, we estimated ${}_1q_0$ and ${}_5q_0$ during three consecutive five-year periods before each survey (i.e., 0-4, 5-9, and 10-14 years before the survey), using a synthetic cohort method (Somoza, 1980; Rustein, 1984). Finally, using UNHCR refugee camp data, we calculated age-specific mortality rates (ASMR) for 0 and 1-4 years of age for each of the 60 camps in year 2008, by the total number deaths in year 2008 in each age group divided by the age-specific mid-year population. We then converted the rates to probabilities of dying (${}_1q_0$ and ${}_4q_1$, respectively), using standard life table methods (Preston 2000), and calculated ${}_5q_0$ as below:

$${}_5q_0 = 1 - \{(1- {}_1q_0)*(1- {}_4q_1)\} \qquad (1)$$

Infant mortality fraction (IMF) refers to the ratio of the two probabilities at given time t:

$$IMF_t = \frac{{}_1q_0(t)}{{}_5q_0(t)}. \qquad (2)$$

Overtime, both infant and under five mortality have reduced but at a different pace, and the ratio captures the disproportional reductions over time in a clear way.

b. Identification of crisis populations in HMD

Using single year data from HMD, we identified two types of crisis populations – Spanish flu pandemic in the early 20th century and World War II, based on excess mortality as compared with mortality in adjacent years. We first examined average percent deviations in ${}_5q_0$ from the non-crisis trends of ${}_5q_0$ over time – i.e., trends calculated by looking at three years before the deviation (flu or war) and three years after – to calculate the expected average values for the year/s of crisis. Let t be the year/s of crisis, then the percentage deviation is calculated as:

$$\text{Percentage deviation} = \left[1 - \frac{R_t}{R_A}\right]*100, \qquad (3)$$

where $R_A = \frac{\sum_{i \neq t} R_i}{6}$ is the average years around the crisis year. Table 1 presents countries and years identified as crisis populations.

Table 1. Countries and years selected to study crisis deviations in infant mortality fractions.

Countries/regions	Years of flu	Years of War
Denmark	1918-1919	1943-1945
England & Wales	1918	1940-1941
France	1918-1920	1943-1945
Finland	1918-1919	1944
Netherlands	1918-1919	1943-1944
Norway	1918-1919	1941-1944
Sweden	1918-1919	1943-1945

c. Data quality in DHS birth histories

Reporting errors in DHS birth histories are usual concerns in examining levels and age-patterns of mortality. In particular, digit preference (i.e., heaping) in reporting age at death on month 12 can underestimate IMF. In order to assess overall data quality, we calculated simple heaping indices at day 7 and month 12 – two ages with most prevalent digit preference – as below:

$$HI(d7) = D(d7) / [\{D(d5) + D(d6) + D(d7) + D(d8) + D(d9)\}/5], \text{ and} \quad (4)$$

$$HI(m12) = D(m12) / [\{D(m10) + D(m11) + D(m12) + D(m13) + D(m14)\}/5], \quad (5)$$

where HI refers to a heaping index, D is the number of deaths reported at certain age in parenthesis, d is day, and m is month. Digit preference on day 28, though suggested to be also common and problematic, cannot be assessed because the unit used in DHS surveys to report age at death changes at this point from days to months. The indices should be close to 1.0 in the absence of error, with an assumption that the number of deaths by day/month is a linear function in the absence of digit preference. Exploratory analyses of the heaping indices showed overall a higher level of heaping on month 12 than on day 7 and no significant associations between the two indices. Survey-period specific observations were categorized into two groups based on levels of the two indices: little or moderate heaping (i.e., $HI(d7) \leq 2.0$ ***and*** $HI(m12) \leq 2.5$) *vs.* severe heaping (i.e., $HI(d7) > 2.0$ ***or*** $HI(m12) > 2.5$).

Table 2. Distribution of data quality among DHS survey-period-specific observations by region (n=469)

Region	Heaping		Total
	Serious	Little or modest	
CEE/CIS	11	3	14
East Asia and Pacific	36	0	36
Latin America and Caribbean	47	45	92
Middle East and North Africa	47	1	48
South Asia	18	21	39
Sub-Saharan Africa	210	30	240
Total	369	100	469

Table 2 above presents distribution of heaping by United Nations regional classification: Central Europe, East Asia and Pacific, Latin American and Caribbean, Middle East and North Africa, South Asia, and sub-Saharan Africa.

2. Analysis Plan

We examined the association between IMF and ${}_5q_0$ on a logarithmic scale, using measurements for consecutive five-year periods from HMD data prior to 1950. We obtained fitted values with 95% confidence intervals from this association, using log linear regression models, and used the model relationship as the non-crisis pattern of IMF relative to ${}_5q_0$. Then we compared patterns of IMF relative to ${}_5q_0$ from various sources of data: HMD crisis population, DHS, and UNHCR refugee camp data. We first compared the associations from the two sets of HMD crisis populations to the HMD non-crisis pattern. We also calculated percent deviation in IMF during crisis years, irrespective of the level of ${}_5q_0$, compared to IMF during three adjacent years before and after the crisis years, as below:

Percentage deviation= $\left[1-\frac{IMF_t}{IMF_A}\right]*100$,

where t is the crisis year(s), and $IMF_A=\frac{\sum_{i\neq t} IMF_i}{6}$ is the average years around the disaster year(s). Finally, we compared associations from DHS (non-crisis patterns in contemporary developing countries) and UNHCR data (patterns from selected crisis populations) to the HMD non-crisis pattern. Further analyses will be focused on examination and understanding ofdifferential patterns by region and data quality, using DHS and UNHCR data.

D. PRELIMINARY RESULTS AND NEXT STEPS

Using HMD data prior to 1950, Figure 1 shows, in Western Europe, an inverse linear relationship between the infant mortality fraction and 5q0. Using the HMD non-crisis pattern as the reference, Figure 2 shows that in non-Sub-Saharan African countries, DHS data do follow the HMD non-crisis pattern. In Sub-Saharan Africa, however, DHS observations typically fall below the HMD non-crisis pattern, indicating a lower proportion of infant deaths at a given level of 5q0. Further region-specific analyses by data quality suggested that the associations do not vary by level of heaping, within each region (Figure 3). Sub-regional analyses, using only Sub-Saharan African countries, suggested that there is no differential pattern between north-east and central-west countries (Figure 4), indicating HIV/AIDS epidemic does not likely explain this departure from the HMD non-crisis pattern.

Examining the UNHCR camp data, monthly data were available for 84 camps in 2008; of these a total of 57 camps had data for the complete year and, of these, 3 camps had no under-five deaths reported for the entire year. Data from the remaining 54 camps were broken down into 36 Sub-Saharan African camps (SSA in the tables and figures) and 18 camps in Asia (referred to as non-SSA in the tables and figures). OF the 54 camps, the average mid-year under-five population in the camps was 3,757 [SD3207] with a range of 534 to 14,019. The mean infant mortality rate was 18.9/1,000 live births [SD13.3] and the under-five mortality rate was 33.9/1,000 live births [SD27.5], substantially lower than mortality rates observed at the national level using DHS. In the scatterplot of the IMF and U5MR presented (Figure 5), we have limited the analysis to 32 camps where the under-five population is >=2000. All camps included in the analysis showed evidence of a lower infant mortality fraction compared to the HMD non-crisis patterns, indicating a lower proportion of infant deaths for a given level of 5q0. In the Sub-Saharan African camps, 5q0 ranges from .05 to .02 while the infant mortality fraction is anywhere from 0.8 to 0.4. While the Asian camp observations are fewer in number, the infant mortality fraction ranges from 0.8 to 0.6 (with one outlier).

The scattered distribution is suggestive of possible migration selection factors, under-reporting of infant deaths due to age-heaping or coverage, or other factors. Exploration of these factors will be undertaken as a next step in data analysis. In order to better understand the populations measured in the UNHCR data, we will assess age-and-sex distributions of populations and deaths and examine other camp-based factors, including age of the camp, migration patterns, data collection systems, and other health indicators.

F. REFERENCES

Checchi F, Roberts L. Documenting mortality in crises: what keeps us from doing better? *PLoS Med* 2008; 5(7): e146.

Checchi F, Roberts L. Interpreting and using mortality data in humanitarian emergencies; A primer for non-epidemiologists. *Humanitarian Practice Network Paper* 2005.

Hill K, Choi Y. Neonatal Mortality in the Developing World. *Demographic Research.* 2006;14:429-452.

Prudhon C, Spiegel PB. A review of methodology and analysis of nutrition and mortality surveys conducted in humanitarian emergencies from October 1993 to April 2004. *Emerging Themes in Epidemiology* 2007, 4:10.

Rustein, S. Infant and child mortality: levels, trends, and demographic differentials. 43. 1984. Voorberg, Netherlands, International Statistical Institute. WFS Comparative Studies.

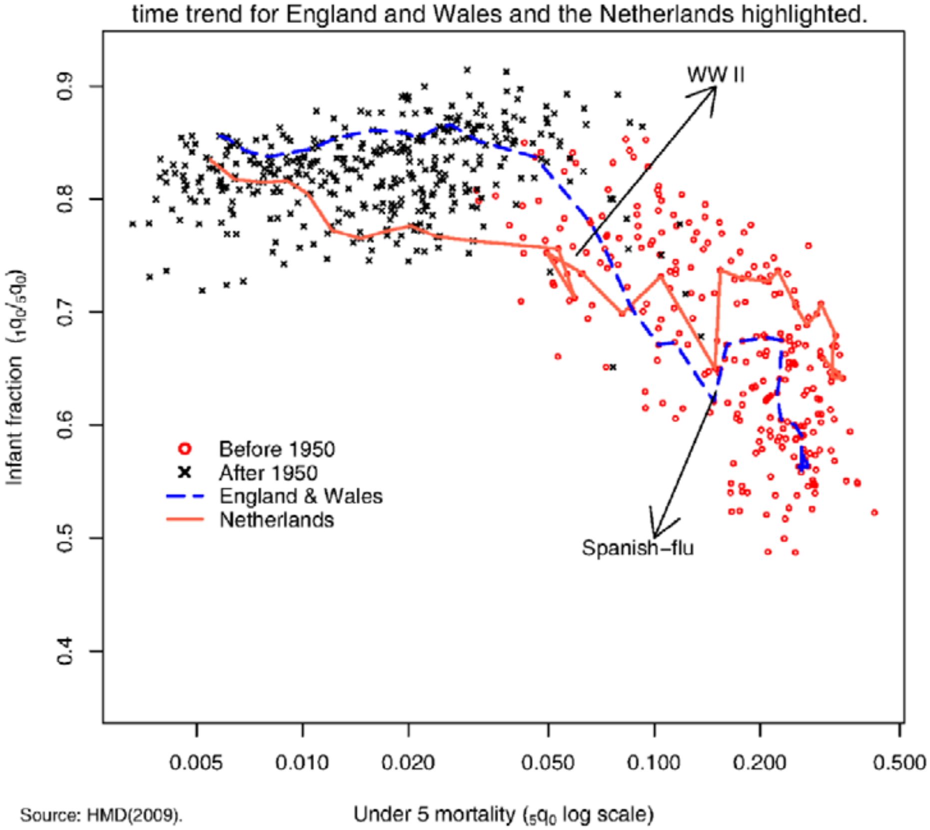
Somoza, J. L. Illustrative analysis: Infant and child mortality in Colombia. 38. 1980. Voorberg, Netherlands, International Statistical Institute. WFS Scientific Reports.

Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, AbouZahr C et al. A scandal of invisibility: making everyone count by counting everyone. *Lancet* 2007; 370: 1569-1577

United Nations Refugee Agency (UNHCR). *Health Information System, Quarterly Bulletin* 2008; 3. (UNHCR; http://www.unhcr.org/protect/4614bce84.html accessed on 1 June 2009).

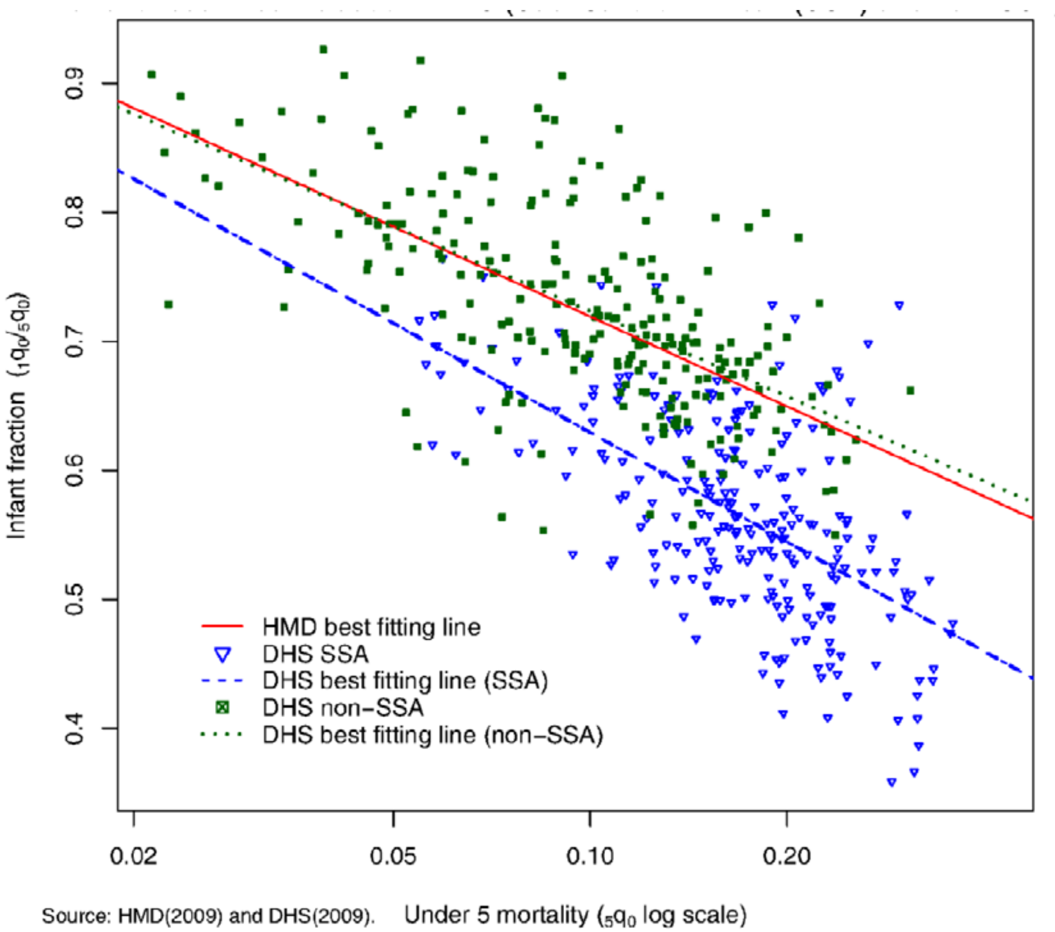
Working Group for Mortality Estimation in Emergencies. Wanted: studies on mortality estimation methods for humanitarian emergencies, suggestions for future research. *Emerging Themes in Epidemiology* 2007, 4:9.

Figure 1. Infant fraction, probability of dying before age one ($_1q_0$) over probability of dying before age five, ($_5q_0$), by levels of $_5q_0$, for industralized countries, time trend for England and Wales and the Netherlands highlighted.



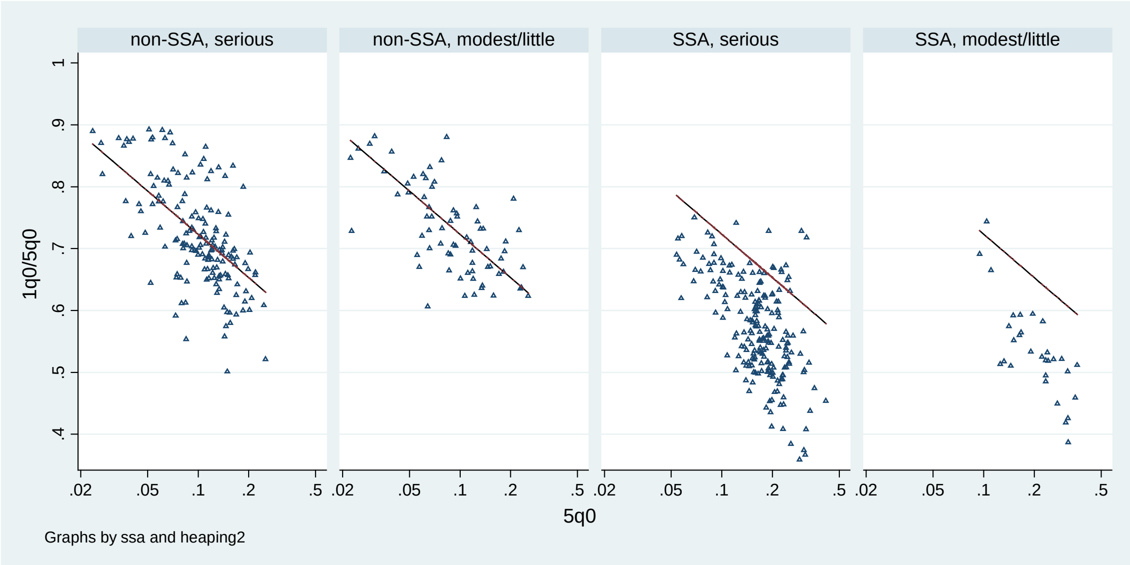
Source: HMD(2009).

Figure 2. Infant mortality fraction by under-five mortality rate in DHS countries by region, comparison to the HMD non-crisis pattern



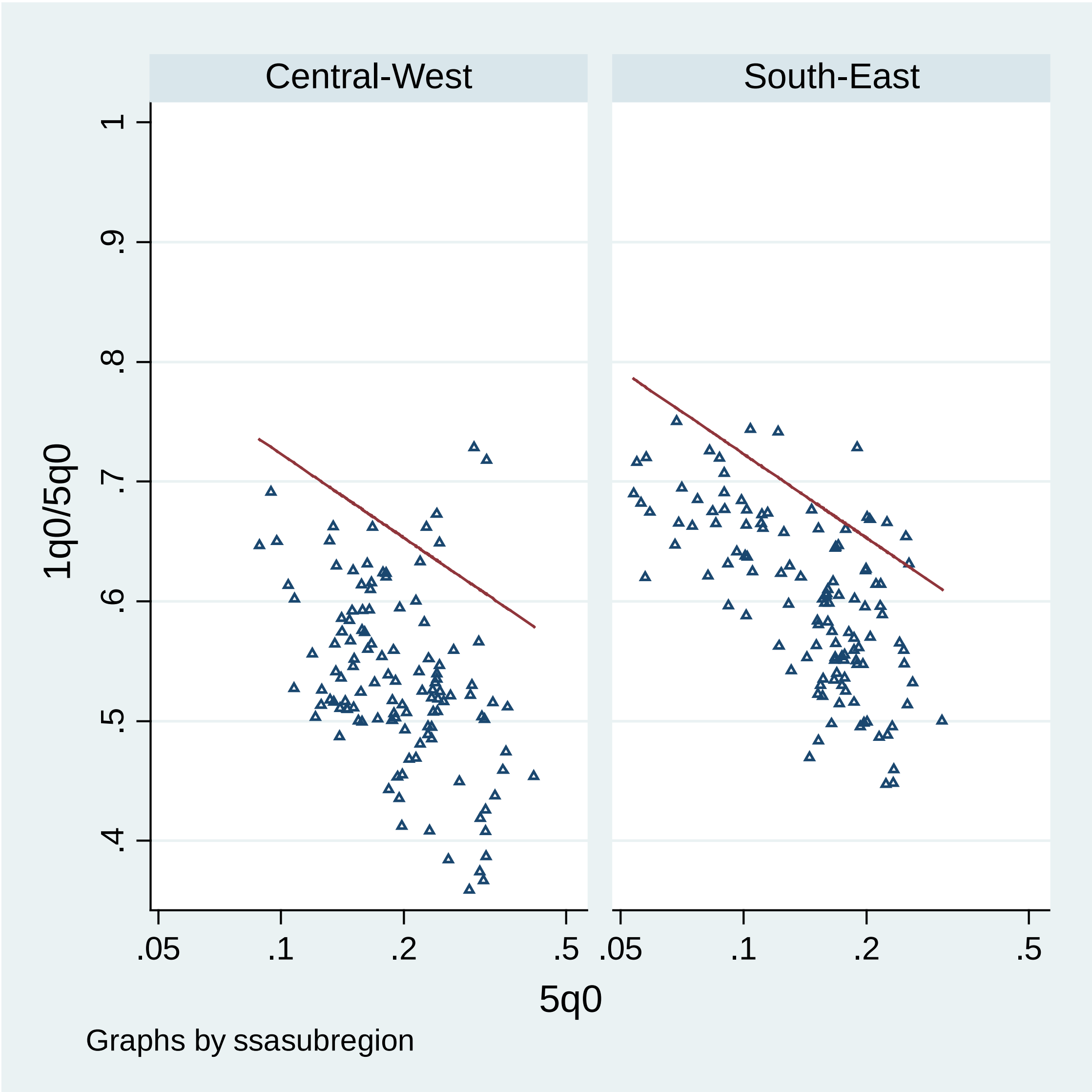
Source: HMD(2009) and DHS(2009).

Figure 3. Infant mortality fraction by level of probability of dying between 0-59 months: DHS survey-period specific observations by region and level of heaping (n=469)



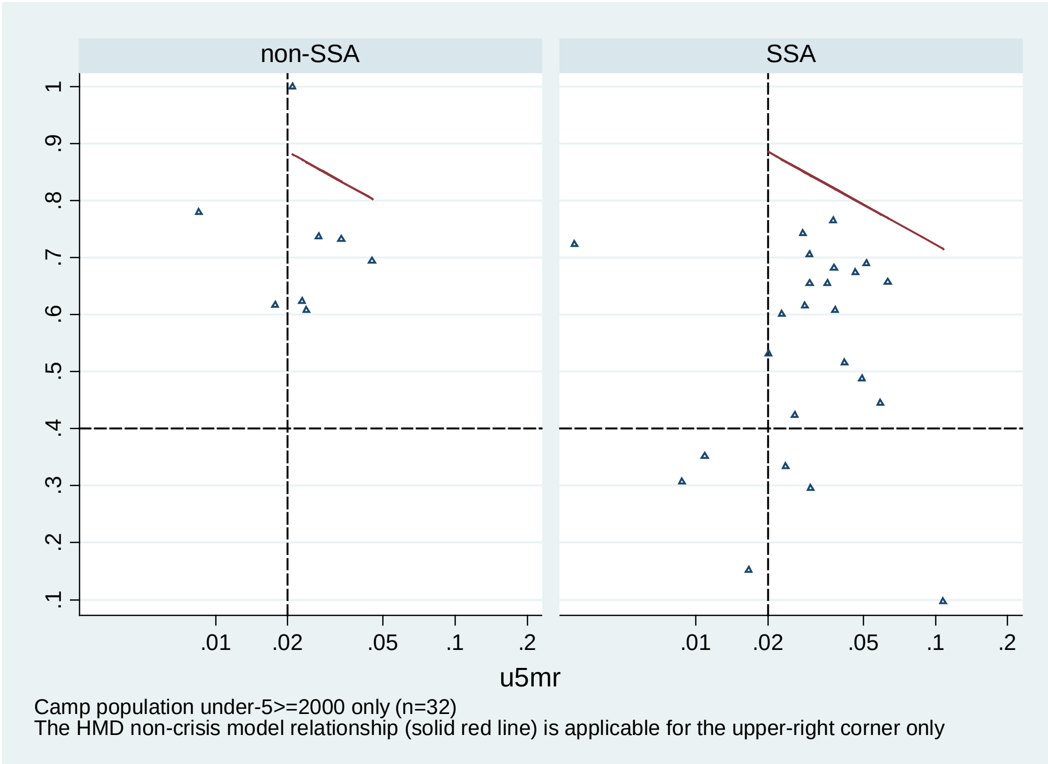
Solid red line is the HMD non-crisis model relationship

Figure 4. Infant mortality fraction by level of probability of dying between 0-59 months: DHS survey-period-specific observations from sub-Saharan African countries by sub-region (n=240)



Solid red line is the HMD non-crisis model relationship

Figure 5. Infant mortality fraction by level of probability of dying between 0-59 months: UNHCR camp observations by region: SSA vs. non-SSA



SSA: sub-Saharan Africa