Age Heaping in Probability Surveys in Low- and Middle-Income Countries: Frequency and Consequences for Mortality Estimation

Kyle Suelflow, Taylor Okonek

August 23, 2024

Abstract

Estimates of under-5 mortality rates (U5MR) are an important indicator of the health of a country. In low- and middle-income countries (LMICs), estimates of U5MR primarily come from probability surveys as opposed to vital registration systems or censuses. The most prevalent of such surveys are the Demographic and Health Surveys (DHS), which are known to have certain quality concerns endemic to probability surveys. One of these quality concerns is age heaping, where a disproportionate number of child deaths are reported at rounded ages, such as 12 months. The exact frequency of age-heaping in DHS surveys is not well-studied, and its impacts on downstream statistical analyses—in particular, the estimation of summary measures of mortality—has not been rigorously explored. In this paper, we will investigate the prevalence of age heaping in DHS surveys, conduct simulations with various "severities" of age heaping, and assess the impact of age heaping on commonly used statistical methods for estimating child mortality summary measures in settings akin to those found in LMICs.

1 Introduction

Summary measures of child mortality like neonatal mortality rates (NMR, the probability of death between zero and one months), infant mortality rates (IMR, the probability of death between zero and 12 months), and under-five mortality rates (U5MR, the probability of death between zero and 60 months) are crucial indicators of the well-being of a country. In high income countries, these summary measures are estimated from civil registration and vital statistics systems (CRVS). In contrast, low- and middle-income countries (LMICs) have poor registration systems (AbouZahr et al., 2015; Yokobori et al., 2021; Mikkelsen et al., 2015). Although legal frameworks for CRVS exist in LMICs, the law itself is written poorly, and enforcement of the law is lacking. Additionally, people may lack knowledge about the benefits of a CRVS, and thus have no real incentive to participate in it (Yokobori et al., 2021; Mikkelsen et al., 2015). In LMICs, births often occur at home and do not get registered in CRVS (Helleringer et al., 2020). Instead of CRVS, probabilistic surveys are used to estimate summary measures of mortality where information is collected via interviews of women aged 15-49 in a representative sample of the population.

The most common of such surveys are the Demographic and Health Surveys (DHS), which collect full birth histories (FBH). In FBH data, women are asked to detail all of their live births: giving the date of birth for each child, indicating whether the child is still alive, and for each deceased child, reporting the child's age of death (Pullum and Becker, 2014). It is important to note that FBH data is not the same as a pregnancy history, where women are asked to detail all pregnancies that they recall, regardless of whether there was a live birth (Pullum and Becker, 2014; AbouZahr et al., 2015). Thus, FBH data does not include stillbirths, miscarriages, or abortions.

DHS surveys have several known issues regarding data quality. One of the most widespread is referred to as age heaping. Age heaping alludes to artificially high death counts at round ages. These ages consist mainly of seven and 28 days, and three, six, 12, and 18 months (Romero Prieto et al., 2021; Hill and Choi, 2006; Pullum and Becker, 2014). This phenomena not only affects estimates of continuous survival curves (Guillot et al., 2022; Okonek et al., 2023), but also estimates of summary measures of mortality. IMR in particular can be underestimated when age heaping occurs due to an upward transfer of deaths. Some deaths that occurred before 12 months are heaped at 12 months, and because IMR does not include deaths at exactly 12 months, this can lead to an underestimation of IMR (Romero Prieto et al., 2021).

Age heaping in the context of child mortality estimation was initially discussed by Goldman et al. (1979), who looked at heaping in Nepal from World Fertility Surveys (WFS). Figure 11 in Goldman et al. (1979) shows significant heaping at ages divisible by six or 12 months in the child mortality period. Large spikes are clearly visible at these ages, most notably at 12 months. Interestingly, heaping at ages ending in zero and five years for adult women is prevalent as well, with large spikes visible on these round ages. Other measures like marital duration and breast feeding duration show similar spikes at round ages.

Curtis (1995) illustrates how age heaping can affect the estimation of child mortality measures. They develop an age heaping index that is used to help determine the prevalence of age heaping at 12 months in different DHS surveys. The index is calculated by taking the number of deaths at 10, 11, 13, and 14 months and averaging them. If deaths follow a linear pattern, then the resulting average should be roughly equal to the number of deaths at 12 months. This average is then divided by the actual number of deaths at 12 months, where an index greater than one indicates the presence of age heaping at 12 months. The surveys they analyzed come from DHS I and DHS II. These are the first two phases of the DHS program. Each subsequent phase builds upon the last, using the latest techniquesCurtis (1995) found that the median index amongst DHS II surveys was five, which is far less than the median index of 11 found by Sullivan et al. (1990) in their assessment of DHS I surveys. To account for age heaping at 12 months, Curtis (1995) adjusts estimates of IMR and child mortality rate (CMR, the probability of death between 12 and 60 months) by shifting 25 percent of the "excess" deaths at 12 months to the period 6-11 months. There is no real reasoning behind the use of 25 percent in this context.

Hill and Choi (2006) investigated how age heaping affects estimates of NMR in children. They focused on heaping at seven days, which is an important age to consider when estimating early NMR, where early NMR is defined as the probability of death before seven days. Deaths heaped at seven days are therefore not considered in the early NMR calculation, rather they contribute to the late neonatal mortality (the probability of death between seven and 28 days) calculation. This can lead to biased estimates of *both* measures, similar to the interaction between deaths heaped at 12 months, and IMR and CMR. Following Curtis (1995), the same heap index was used, except instead of looking at deaths at 12 months, Hill and Choi (2006) looked at deaths at 7 days. This heap index was calculated by dividing the number of deaths at seven days by 1/5 the deaths from five to nine days. An index of one indicates that there was no evidence of age heaping, as deaths at seven days were equal to the average number of daily deaths in the five to nine day range. Hill and Choi (2006) found strong evidence for the occurrence of age heaping, particularly in sub-Saharan Africa. Forty percent of the 108 DHS surveys included in the study had a heap index greater than 2.5, and in sub-Saharan Africa, 92 percent of surveys exhibited this behavior.

This study aims to be a principal resource for understanding how frequently age heaping occurs in DHS surveys in sub-Saharan Africa, and how existing methods for estimating child mortality are affected by age heaping. The methods for estimating child mortality are compared in terms of their sensitivity to age heaping are the log-quad model (Guillot et al., 2022), the discrete hazards model (Li et al., 2019), and the continuous, parametric pseudo-likelihood approach (Okonek et al., 2023), which are detailed in section 3. While this is not an exhaustive list of methods used to estimate child mortality in LMICs, they represent a range of recently developed approaches that are either commonly used or are particularly equipped to address issues of age heaping. We examine the effects age heaping has on these methods through a simulation study. By using simulated data rather than data from DHS, we can control aspects of data quality, and in particular, the amount of age heaping that occurs. Importantly, this allows for isolation the effects of age heaping. In DHS surveys, there are almost certainly data quality issues in addition to age heaping, namely issues of mis-classification of neonatal deaths Helleringer et al. (2020), and potential survivor bias Hill and Choi (2006). These quality concerns are outside the scope of this paper, though this simulation framework could be extended to further explore these issues. In this paper, we examine how and why age heaping occurs, and when it occurs in the age schedule (age schedule refers to the range of ages from zero to 60 months). We give examples of specific DHS surveys where age heaping occurs to illustrate the challenges it poses for estimating child mortality. We then describe the methods that are compared in our simulation Guillot et al. (2022), Li et al. (2019), Okonek et al. (2023). Finally, we detail our simulation set-up, with justification for the values of the simulation parameters that vary. The results of the simulation study are given in section 5. All of the code can be found via https://github.com/ksuelflo/age_heaping2024.

2 Age Heaping

Age heaping in DHS surveys is a barrier to accurate estimates of a number of important health measures. (Pullum and Staveteig, 2017) investigated the prevalence of age heaping in the context of interviewed women, and whether their was preference given towards ages that ended in zero or five. We investigate the prevalence of age heaping in the context of child mortality, where heaping at 12 months is the most common age. Every DHS survey from sub-Saharan Africa beginning in 2005 or later was included in this study. We used a “heap index” to compare prevalence across surveys. This index is similar to the one used in (Hill and Choi, 2006; Curtis, 1995). The heap index is calculated by dividing the number of deaths at 12 months by 1/5 the total number of deaths between 10 and 14 months.

3 Methods

3.1 Guillot

The log-quad model which (Guillot et al., 2022) proposes is relatively unique in its ability to predict a detailed age schedule of mortality, with 22 different age intervals at which mortality rates are predicted. Their predictions are based off of the under-five mortality database (U5MD), which is comprised of high quality CRVS data. They take a conservative approach to building the database, which is important to consider when evaluating the model predictions. Datasets with concerns over data quality are not apart of the database. This can lead to biased estimates of U5MR, IMR, or NMR in places which exhibit different characteristics than the U5MD.

$$\ln\left[q(x)\right] = a_x + b_x \cdot \ln\left[q(5y)\right] + c_x \cdot \ln\left[q(5y)\right]^2 + v_x \cdot k. \tag{1}$$

The model takes in the overall level of U5MR, and up to 21 $q(x)$ values, where $q(x)$ is defined as the probability of death before age x. $\{a_x, b_x, c_x, v_x\}$ are a set of age specific coefficients. Each x corresponds to one of the 22 age intervals mentioned above. The model predicts k, which is a shape parameter. When k is zero, the model predicts an age schedule of mortality that is an average of the U5MD. When k is greater than zero, the model predicts an "early" mortality schedule, where neonatal mortality is higher than average. Values of k less than zero indicate a "late" mortality schedule, characterized by relatively high levels of mortality later in the age schedule. However, we know that some LMICs, in particular sub-Saharan Africa countries, exhibit both "early" and "late" behaviour (Eilerts et al., 2021; Romero Prieto et al., 2021; Guillot et al., 2022; Verhulst et al., 2022).

3.2 Mercer

Mercer et al. (2015) utilizes a discrete-hazards approach to estimating child mortality. They calculate U5MR as follows:

$$_5q_0 = 1 - \prod_{j=1}^{J}(1 -_{n_j} q_{xj}) \tag{2}$$

The first 60 months are divided into $J = 6$ intervals: [0, 1),[1, 12),[12, 24),[24, 36),[36, 48),[48, 60) with $(x_1,...,x_6) = (0, 1, 12, 24, 36, 48)$ and $(n_1,...,n_6) = (1, 11, 12, 12, 12, 12)$ (Mercer et al., 2015). It is important to note that mortality is assumed constant within these intervals. Logistic regression is used to estimate the probability of dying conditional on the state of the child at the beginning of the interval.

3.3 Pseudo-Likelihood

4 Simulation

We varied seven different parameters, shown in 2, to examine the impact they might have on summary measures of child mortality obtained via the methods described in section 3. The lognormal mean and standard deviation parameter values were obtained as a result of fitting a lognormal curve to a selection of different DHS surveys using the R package *pssst* Okonek et al. (2023). The lognormal PDF is shown below:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{1}{2\sigma^2}(\log(x)-\mu)^2} \tag{3}$$

Lognormal Parameters			
Country	Year	μ	$\log(\sigma)$
Namibia	2006	16.69752	2.107787
Ghana	2008	17.06582	2.192750
Madagascar	2008	16.34247	2.112932
Nigeria	2008	11.23414	1.873644
Burkina Faso	2010	11.44673	1.811881
Malawi	2010	13.92794	2.020223
Rwanda	2010	14.95992	2.047975
Cameroon	2011	12.01134	1.858946
Namibia	2013	19.63663	2.247907
Chad	2014	11.46442	1.842216
Ghana	2014	19.95666	2.333236
Rwanda	2014	18.54880	2.170704
Malawi	2015	18.46945	2.255041
Cameroon	2018	16.15746	2.128927
Nigeria	2018	12.87478	1.994495
Mauritania	2019	23.62952	2.419955
Rwanda	2019	22.16235	2.377468
Burkina Faso	2021	19.01027	2.176021
Madagascar	2021	16.77247	2.151735
Ghana	2022	21.12636	2.277469

Table 1: Year refers to the year the survey was conducted. One five year period ending on the year the survey was conducted was used to fit the lognormal curve. Parameters used to fit a lognormal curve via the `pssst` R package (Okonek et al., 2023).

The lognormal distribution uses parameters σ and μ. We randomly chose 10 different sub-Saharan Africa countries, and fit lognormal curves to all DHS surveys with a survey year greater than 2005. Based on the resulting values shown in table 1, we chose 4 different combinations of lognormal parameters, μ_i and σ_j, where $\mu = \{12, 15, 15, 20\}$ and $\log(\sigma) = \{1.9, 2.1, 2.3, 2.3\}$. We compare sample sizes of 100 and 500 children born at each month, respectively. The only age we considered for age heaping was at 12 months, as 12 months is the most commonly referenced age in the literature (Hill and Choi, 2006; Espeut and Becker, 2015; Helleringer et al., 2020; Romero Prieto et al., 2021; Guillot et al., 2022). Heaping at three, six, or 18 months, as Romero Prieto et al. (2021) outlines, is an area for further exploration. The proportions of deaths chosen to be heaped at 12 months, shown in the "Proportion heaped" row in table 2, offer a range to evaluate model performance in the presence of various severities of age heaping. For example, when the proportion heaped is 20%, deaths that fall in the specified range ([9,21], for example) have a 20% chance of being heaped at 12 months. The range parameter is a range of months, where a death that falls within the range has a chance to be heaped depending upon the proportion parameter. Guillot et al. (2022) suggested excluding ages from nine to 21 months due to the possibility of age heaping, so we propose this range as one option. The eight to 24 month range comes from Romero Prieto et al. (2021), where they defined ages that are affected by heaping at 12 months as the eight to 24 month range. A symmetric range of six to 18 months is another range that was included. For every simulation there are five periods of 60 months each.

Simulation Settings	
Parameter	Values
Lognormal mean	12, 15, 20
Lognormal standard deviation	1.9, 2.1, 2.2
Sample size	100, 500
Age of heaping	12 months
Proportion heaped	10%, 20%, 50%
Range	[9,21],[8,24],[6,18]
Period Length	60 months

Table 2: Sample size refers to the number of children born at the beginning of each month.

5 Results

6 Discussion

References

AbouZahr, C., De Savigny, D., Mikkelsen, L., Setel, P. W., Lozano, R., and Lopez, A. D. (2015). Towards universal civil registration and vital statistics systems: the time is now. *The Lancet*, 386(10001):1407–1418.

Curtis, S. L. (1995). *Assessment of the quality of data used for direct estimation of infant and child mortality in DHS-II surveys.* Macro International Incorporated.

Eilerts, H., Prieto, J. R., Eaton, J. W., and Reniers, G. (2021). Age patterns of under-5 mortality in sub-saharan africa during 1990–2018: a comparison of estimates from demographic surveillance with full birth histories and the historic record. *Demographic research*, 44:415.

Espeut, D. and Becker, S. (2015). The validity of birth and pregnancy histories in rural bangladesh. *Journal of Health, Population and Nutrition*, 33:1–11.

Goldman, N., Coale, A. J., and Weinstein, M. (1979). The quality of data in the nepal fertility survey. International Statistical Institute.

Guillot, M., Romero Prieto, J., Verhulst, A., and Gerland, P. (2022). Modeling age patterns of under-5 mortality: results from a log-quadratic model applied to high-quality vital registration data. *Demography*, 59(1):321–347.

Helleringer, S., Liu, L., Chu, Y., Rodrigues, A., and Fisker, A. B. (2020). Biases in survey estimates of neonatal mortality: results from a validation study in urban areas of guinea-bissau. *Demography*, 57(5):1705–1726.

Hill, K. and Choi, Y. (2006). Neonatal mortality in the developing world. *Demographic research*, 14:429–452.

Li, Z., Hsiao, Y., Godwin, J., Martin, B. D., Wakefield, J., Clark, S. J., with support from the United Nations Inter-agency Group for Child Mortality Estimation, and its technical advisory group (2019). Changes in the spatial distribution of the under-five mortality rate: Small-area analysis of 122 DHS surveys in 262 subregions of 35 countries in Africa. *PloS One*, 14(1):e0210645.

Mercer, L. D., Wakefield, J., Pantazis, A., Lutambi, A. M., Masanja, H., and Clark, S. (2015). Space-time smoothing of complex survey data: small area estimation for child mortality. *The annals of applied statistics*, 9(4):1889.

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., De Savigny, D., Lozano, R., and Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *The Lancet*, 386(10001):1395–1406.

Okonek, T., Wilson, K., and Wakefield, J. (2023). A pseudo-likelihood approach to under-5 mortality estimation.

Pullum, T. and Staveteig, S. (2017). *An Assessment of the Quality and Consistency of Age and Date Reporting in DHS Surveys, 2000-2015*. ICF International.

Pullum, T. W. and Becker, S. (2014). *Evidence of omission and displacement in DHS birth histories*. ICF International.

Romero Prieto, J., Verhulst, A., and Guillot, M. (2021). Estimating the infant mortality rate from dhs birth histories in the presence of age heaping. *PLoS One*, 16(11):e0259304.

Sullivan, J. M., Bicego, G. T., and Rutstein, S. O. (1990). Assessment of the quality of data used for the direct estimation of infant and child mortality in the demographic and health surveys.

Verhulst, A., Prieto, J. R., Alam, N., Eilerts-Spinelli, H., Erchick, D. J., Gerland, P., Katz, J., Lankoande, B., Liu, L., Pison, G., et al. (2022). Divergent age patterns of under-5 mortality in south asia and sub-saharan africa: a modelling study. *The Lancet global health*, 10(11):e1566–e1574.

Yokobori, Y., Obara, H., Sugiura, Y., and Kitamura, T. (2021). Gaps in the civil registration and vital statistics systems of low-and middle-income countries and the health sector's role in improving the situation. *Global Health & Medicine*, 3(4):243–245.