Spatial and Spatio-temporal Epidemiology 11 (2014) 89–107

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste

Estimation of district-level under-5 mortality in Zambia using birth history data, 1980–2010

Laura Dwyer-Lindgren [a,*], Frank Kakungu [b], Peter Hangoma [c], Marie Ng [a], Haidong Wang [a], Abraham D. Flaxman [a], Felix Masiye [c], Emmanuela Gakidou [a]

[a] Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA
[b] Central Statistical Office of Zambia, Lusaka, Zambia
[c] Department of Economics, University of Zambia, Lusaka, Zambia

ARTICLE INFO

Article history:
Received 21 February 2014
Revised 16 September 2014
Accepted 18 September 2014
Available online 13 October 2014

Keywords:
Small area estimation
Geospatial analysis
Spatial–temporal smoothing
Birth history methods
Under-5 mortality
Child mortality

ABSTRACT

Birth history data—the primary source of data on under-5 mortality in developing countries—are infrequently used for subnational estimates due to concerns over small sample sizes. In this study we consider different methods for analyzing birth history data in combination with various small area models. We construct a simulation environment to assess the performance of different combinations of birth history methods and small area models in terms of bias, efficiency, and coverage. We find that performance is highly dependent on the birth history method applied and how temporal trends are accounted for. We estimated trends in district-level under-5 mortality in Zambia from 1980 to 2010 using the best-performing model. We find that under-5 mortality is highly variable within Zambia: there was a 1.8-fold difference between the lowest and highest levels in 2010, and declines over the period 1980 to 2010 ranged from less than 5% to more than 50%.

1. Introduction

Under-5 mortality—the probability that a child will die before reaching the age of 5 if he or she experiences the age-specific risks of death observed in the current year—is a basic and widely used indicator of child health and survival. However many countries, particularly those where under-5 mortality is high, lack registration systems to record deaths, complicating measuring and tracking trends in this indicator. In these countries, what knowledge we do have of under-5 mortality is based on survey data, primarily in the form of birth histories where women are interviewed about the mortality experience of their children. While birth history data have been widely used for estimating under-5 mortality at a national level, there are relatively few instances (Bangha and Simelane, 2007; Storeygard et al., 2008; Singh et al., 2011; Bauze et al., 2012) where they have been used to estimate mortality at a fine subnational level, due primarily to concerns about small sample sizes. As a consequence, knowledge about subnational trends in under-5 mortality is mostly lacking in the very countries where under-5 mortality is the greatest threat.

Small area models are statistical models which address the issues raised by small sample sizes by explicitly accounting for the large sampling variance and exploiting spatial and temporal relatedness to increase predictive power. These types of models have been occasionally applied to birth history data on mortality in the past, but the focus has usually been on the relationship between

* Corresponding author. Tel.: +1 206 897 2823.
E-mail addresses: ladwyer@uw.edu (L. Dwyer-Lindgren), FKakungu@zamstats.gov.zm (F. Kakungu), peterhangoma555@yahoo.com (P. Hangoma), marieng@uw.edu (M. Ng), haidong@uw.edu (H. Wang), abie@uw.edu (A.D. Flaxman), fmasiye@yahoo.com (F. Masiye), gakidou@uw.edu (E. Gakidou).

http://dx.doi.org/10.1016/j.sste.2014.09.002
1877-5845/

socioeconomic or environmental factors and the risk of death, not prediction of under-5 mortality or other similar childhood mortality indicators (Gemperli et al., 2004; Adebayo et al., 2004; Kandala and Ghilagaber, 2006; Kazembe et al., 2007; Kazembe and Mpeketula, 2010; Asiimwe et al., 2011; Chin et al., 2011). This is likely because mortality data pose several challenges in the context of small area methods. Most small area models are designed for count data; under-5 mortality is a complex construct, however, not a simple count. Further, in order to derive estimates of under-5 mortality from certain types of birth history data, demographic models must be employed, adding an additional modeling step. Finally, there are a number of different methods available for analyzing birth history data and it is not obvious which is the best suited for combination with small area methods.

In this analysis we explore different ways of combining birth history methods and small area models to estimate under-5 mortality at a subnational level from multiple data sources. Specifically, we use simulation to construct a validation environment in which to test various combinations of birth history methods and small area models and to select a best method. We then apply this method to birth history data in Zambia in order to estimate under-5 mortality from 1980 to 2010 at the district level. The paper is organized as follows: in Sections 2 and 3, respectively, we outline the birth history methods and small area models considered. We then describe how we constructed a simulation environment for testing the various combinations of birth history methods and small area models in Section 4. In Section 5 we compare the performance of all methods considered and select a best method. We describe an application of this method to data available in Zambia in Section 6. Finally we discuss the findings and implications of this research in Section 7.

2. Birth history methods

There are two types of birth histories routinely collected which allow for estimation of under-5 mortality over time in the years preceding a survey or census: complete birth histories (CBH) and summary birth histories (SBH). In a CBH, women are asked detailed questions about each child they have given birth to, including the date of birth, survival status, and (when applicable) age at death. In contrast, in a SBH women are asked only for the total number of children they have given birth to and the number of these children that have died. CBH contain sufficient information to calculate under-5 mortality directly, but SBH require demographic methods which use regression models to relate information about total children born and died to under-5 mortality. SBH are more frequently collected, however, because conventional wisdom holds that they impose a considerably smaller time burden during data collection than CBH.

The methods we used for analyzing CBH data have been described elsewhere (Dwyer-Lindgren et al., 2013a). Briefly, the months that each child lived before death or reaching age 5 alive are divided up by age group (month 0, months 1-11, year 1, year 2, year 3, and year 4) and by time period. Then for each time period, the monthly probability of surviving in each age group is calculated as the ratio of deaths that occurred at that age in that time period to the number of months lived at that age in that time period. Under-5 mortality (denoted $q5$) is then calculated for each time period by subtracting from 1 the product of the monthly probabilities of survival (denoted p):

$$q5_t = 1 - p_{t,mo.0} \cdot (p_{t,mos.1-11})^{11} \cdot (p_{t,yr1})^{12} \cdot (p_{t,yr2})^{12} \cdot (p_{t,yr3})^{12} \cdot (p_{t,yr4})^{12} \quad (1)$$

Although under-5 mortality is calculated directly from CBH data, analysts must still chose how to group data into time periods, with shorter time periods allowing for a more nuanced analysis of time trends but longer periods providing larger sample sizes and more stable estimates (Pedersen and Liu, 2012). In this analysis, we analyzed CBH data using one-, two-, and five-year periods, pooling data across all available sources.

We used the suite of methods proposed by Rajaratnam et al. (2010) and Lozano et al. (2011) to analyze SBH data. These methods use regression models to relate under-5 mortality at various time intervals prior to a survey to quantities available from SBH data, in particular the fraction of children ever born who have died, as well as the mother's age or reported time since first birth. These authors describe four individual methods that make use of different combinations of the data available from a SBH—the maternal age cohort method (MAC), time since first birth cohort method (TFBC), maternal age period method (MAP), and time since first birth period method (TFBP)—as well as a combined method which synthesizes the estimates from the other four methods. The cohort methods group the SBH data by cohorts of women defined either by their age (MAC) or by the time since first birth (TFBC); regression models are then used to generate one estimate of under-5 mortality from the data for each cohort of women. The period methods use empirically derived distributions of births and deaths prior to survey, indexed either by mother's age (MAP) or time since first birth (TFBP), to distribute the reported births and deaths across periods prior to the survey; regression models are then used to generate one estimate for each year prior to survey based on the births and deaths distributed to that year. In this analysis, we analyzed SBH in three ways: using just the cohort methods (MAC and TFBC), using just the period methods (MAP and TFBP), and using the combined method.

It is common for both a SBH and a CBH to be collected in a survey whereas in censuses generally only a SBH is collected. Thus in addition to the choice about methods to apply to CBH and SBH data, analysts must also decide what data, of that available, should be used from a given source. To this end, we considered a total of ten different 'data formats' or combinations of available birth history data and analysis methods. These formats are summarized in Table 1. The first three utilize SBH data only from all available sources. The next three utilize CBH data only and are consequently only applied to surveys. The final four formats combine SBH and CBH data in various ways.