Low-cost sibling methods for measuring adult mortality

Bruno Masquelier[1], Ashira Menashe-Oren[1], Georges Reniers[2], Ian M. Timæus[2,3]
[1] UCLouvain, Belgium [2]London School of Hygiene & Tropical Medicine, UK [3]University of Cape Town, South Africa

Introduction

In many low- and middle-income countries, population-based surveys and censuses remain key data sources for estimating mortality as death registration systems are under-developed [1]. With the inclusion of questions on sibling survival in several survey programs, most notably the Demographic and Health Surveys (DHS), the empirical evidence on adult mortality has greatly expanded since the 1990s.

Sibling survival histories (SSH) can take two forms: *full sibling histories*, with questions on ages at survey or ages at death and the timing of death, and *summary sibling histories*, collecting information only on the total number of (adult) siblings and those who were alive at the time of the survey. Full SSH are widely used to produce global health estimates of adult and pregnancy-related mortality [2, 3], to measure the effectiveness of health programs [4] or to monitor mortality increases associated with violence [5] or HIV/AIDS [6]. Collecting full SSH is time-consuming, however. In a study in Senegal, the median duration of an interview to collect full SSH and basic socio-demographic characteristics was 30 minutes [7]. For this reason, not all survey programs collect SSH, and they are usually collected from women only. Full SSH have never been included in censuses. They are also unsuitable for rapid surveys conducted by mobile phones such as those being considered to track COVID-related excess mortality [8].

The length of the interview is considerably reduced with summary SSH. The duration of exposure to the risk of dying can be approximated based on the age of respondents and reference periods can be estimated based on stable population theory. In this study, we evaluate the performance of an existing indirect method, developed by Timæus and colleagues (2001) [9] and systematically compare these with direct estimates using SSH data from 53 countries. We also apply this indirect method in microsimulations where the "true" mortality rates are known.

An important limitation of this *cohort-derived indirect method* is that estimates refer to a relatively distant past (3 to 15 years before the survey). This method also assumes that trends in mortality have been regular, making it inadequate for estimating mortality in settings disrupted by conflict, disasters or epidemic outbreaks. A third disadvantage is that mortality estimates are based on both recent deaths and those that occurred in a more distant past, which are more likely to be omitted by respondents.

To address these limitations, we introduce a new *period-derived indirect approach*, which focuses solely on recent deaths. This method requires input data on the total number of siblings still alive, those who were still alive five years ago (or another reference period in the recent past), and the total number of siblings ever born. The age of the respondent is used to approximate the age of siblings at the start of the reference period, adjusting for the effect of past trends in mortality on age differences between surviving siblings. As for the cohort-derived indirect method, we evaluate the performance of this estimation approach using existing DHS and test it in microsimulations.

Data and methods

This study is based on the sibling survival data collected in the 146 DHS that included an adult mortality module. The surveys were conducted between 1992 and 2019 in 53 countries.

To evaluate the indirect methods in the absence of reporting errors, we generate fictitious populations via microsimulation, using the SOCSIM model [10, 11]. SOCSIM is a closed model, meaning that new individuals can be introduced into the population only through birth. This allows for the identification of sibships, as all individuals have a mother who is herself identified. The direct and indirect mortality rates obtained from SSH can be evaluated against the underlying mortality rates introduced in the simulations.

We generate 96 different stable populations, corresponding to different levels and age patterns of mortality and fertility, and a constant growth rate. We retain the set of fertility and mortality parameters that Timæus et al. (1992) used to estimate the relationship between mortality and orphanhood (Table 1) [12]. Mortality is represented by relational logit models, specifying 3 values for the α parameter (capturing variations in the level of mortality) and 2 values for the β parameter (capturing differences in age patterns) and using the Brass's general standard. Fertility is modelled with the Brass relational model for fertility and the Booth standard [13], with 2 values of α_f (capturing the age location of the fertility schedule) and 2 values of β_f (capturing the spread of the fertility schedule) [14]. Altogether, these 96 simulations cover a wide range of possible demographic profiles. The corresponding life expectancies at birth range from 35,8 to 72,8 years, while the mean age of the fertility schedule ranges from 25.7 to 33.0 years. The simulations run for 150 years and the final populations reach about 40,000 surviving members.

Mortality		Fertility
α	0.2, -0.2, -0.6, -1.0	-0.5, -0.2,0.1,0.4
β	0.8, 1.1	0.7, 1.0, 1.3

Table 1 – *Parameters used to set up the microsimulations (with a growth rate set at 2%)*

Results

An example of the direct calculation is provided in Table 2 from the full SSH collected in a recent DHS conducted in Sierra Leone in 2019, among females. Age-specific mortality rates (ASMR) can be summarized into ${}_{35}q_{15}$, the risk that a person aged 15 dies before reaching age 50 when facing the ASMR of a specific period. Confidence intervals around the ASMR and ${}_{35}q_{15}$ are obtained using a stratified jackknife approach that accounts for the sample design of DHS. In the 2019 Sierra Leone DHS, the ${}_{35}q_{15}$ probability for females is estimated at 168 ‰(95% CI: 149 - 187) for the period 0-6 completed years before the survey.

The *cohort-derived indirect method* developed by Timæus and colleagues (2001) is applied to the same survey in Table 3 (where ${}_{5}S_{n-5}^{+15}$ refers to the proportions of siblings who survived to the time of data collection, among those who reached their 15th birthday). The value obtained from 25-to 29-year-olds (167 ‰) is identical to the estimate obtained directly for the 0-6 year period before the survey, but the two surrounding estimates are much lower, despite referring to a similar period. In Figure 1, the comparison is extended to all surveys conducted in Sierra Leone, showing relatively good consistency between direct and indirect estimates. Data quality problems are apparent, as the trend inferred from each survey separately shows a gradual increase in mortality, while the overall trend derived from the three successive surveys shows stagnation or a slight decline (see the point estimates published in DHS reports).

Age group	Deaths	PY	ASMR (‰)
15-19	81	28878	2.81
20-24	110	31728	3.48
25-29	118	29997	3.94
30-34	121	25059	4.84
35-39	118	19233	6.15
40-44	87	12122	7.22
45-49	53	6348	8.36
$_{35}q_{15}$	168	(95% CI:149-187)	

Table 2 – *Direct estimates, 2019 Sierra Leone DHS, females, 0-6 years before the survey*

n	Age group	Sisters reaching age 15	Sisters surviving	$_5S_{n-5}^{+15}$	$_{n-15}p_{15}$	Years since survey	$_{35}q_{15}$ (‰)
25	20-24	4123	4020	0.975	0.976	3.3	134.9
30	25-29	5014	4778	0.953	0.947	5.6	166.5
35	30-34	3639	3456	0.950	0.943	7.7	132.5
40	35-39	4443	4130	0.929	0.921	9.6	134.1
45	40-44	2591	2376	0.917	0.910	11.4	118.6
50	45-49	2345	2053	0.876	0.867	13.0	132.8

Table 3 – *Cohort-derived indirect estimates using coefficients from Timæus et al. (2001), 2019 Sierra Leone DHS, females, 0-13 years before the survey*

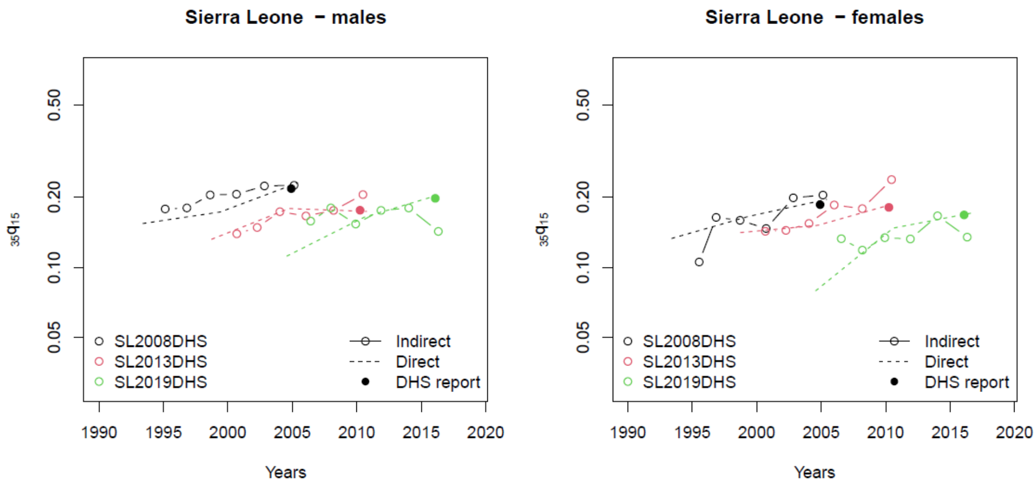

Figure 1 – *Comparison of direct and indirect estimates obtained from sibling survival data in Sierra Leone*

The *period-derived indirect method* introduced in this study requires that three questions are asked: (1) How many of your sisters born to the same mother are currently alive (excluding the respondent)? How many other sisters from the same mother have died in the last 5 years? (3) How many other sisters from the same mother died before that?

The procedure for estimating mortality from these summary SSH involves four steps. First, a set of age-specific probabilities of dying between birth and *t-5* is obtained from the proportions of surviving siblings at *t-5* tabulated by age of respondents [15]. Second, based on stable population theory, a theoretical distribution of age differences between the respondent and her siblings is derived, considering all siblings ever born, and subsequently adjusted so that it refers only to siblings who were still alive at *t*-5. Figure 2 shows the distributions of the age difference between a respondent and her siblings, as computed analytically in stable populations if we use the 96 sets of parameters summarized in Table 1 and assume that the mothers' achieved fertility distributions match that of the population. The distribution of sibling age differences observed in the 2019 DHS in Sierra Leone among respondents aged 15-49 is superimposed in the graph. It is remarkably close to the modelled distributions obtained when β_f, the parameter referring to the dispersion of the age pattern of fertility, is set at 1.3. This parameter can be approximated based on information contained in summary SSH, such as ratios of the age-specific mean number of siblings ever born.

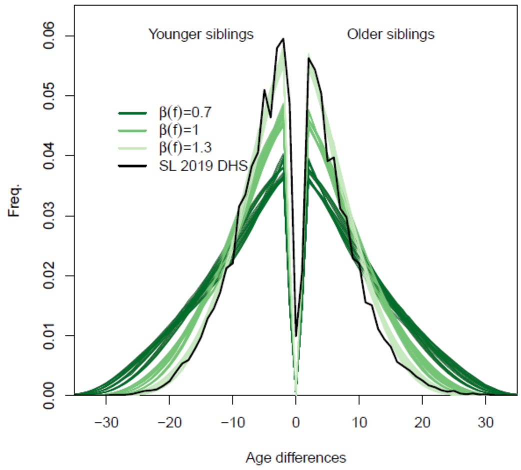

Figure 2 – *Sibling age difference distributions corresponding to 96 stable populations and distribution observed in the 2019 Sierra Leone DHS*

The age of siblings can easily be inferred from the age of respondents when considering siblings ever born, but mortality plays a role when considering those still alive at *t*-5. This is because younger siblings have been exposed to risks of dying during a shorter period than the respondent, while older siblings have been exposed up to an older age. Since we have estimates of mortality before *t*-5 from the first step, the effect of past mortality on the age difference between the respondents and her siblings still alive at *t*-5 can be accounted for. The distribution needs to be revised by multiplying each birth cohort by the survivorship probability up to *t*-5. The distribution is also adjusted for the age group of respondents (e.g. those aged 15-19 have no younger sibling born more than 20 years ago). The third step is to randomly allocate a date of death in the last 5 years, assuming that deaths are uniformly distributed over this recent period. The fourth and final step consists of estimating mortality directly, based on imputed dates of birth and death. Overall, imputing the dates of birth and death has the added benefit of mitigating the effects of age heaping or systematic misstatement of ages at death and at survey.

In the 2019 Sierra Leone DHS, the probability ${}_{35}q_{15}$ estimated from this period-derived method is 158 ‰ (95% CI 134-181), which is slightly higher (by 8%) than the value calculated directly from the same survey for the 5-year period prior to the survey (147 ‰ - 95% CI 126-167). Summary SSH centered on recent deaths also allow the reconstruction of a synthetic cohort for the 5-year period preceding the survey, and the use of the multiplier coefficients of Timæus and colleagues (2001) directly on the proportions of siblings surviving within that cohort. Applied to the 2019 Sierra Leone DHS survey, this approach yields an estimate of the probability ${}_{35}q_{15}$ of 156 ‰, also very close to the direct estimate without the need for imputation.

These preliminary results suggest recent mortality can be measured with good accuracy using summary SSH. The full paper will extend the analysis by applying direct and indirect methods to all DHS surveys and using the various microsimulations to test their robustness.

References

[1] L. Mikkelsen *et al.*, “A global assessment of civil registration and vital statistics systems: monitoring data quality and progress,” *The Lancet*, vol. 386, pp. 1395–1406, 2015.

[2] H. Wang *et al.*, “Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the global burden of disease study 2019,” *The Lancet*, vol. 396, pp. 1160–1203, 2020.

[3] L. Alkema *et al., “*Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group”, *The Lancet*, 387, 462-474, 2016

[4] E. Bendavid, C. Holmes, J. Bhattacharya, and G. Miller, “HIV development assistance and adult mortality in Africa,” *JAMA*, vol. 307, no. 19, pp. 2060–2067, 2012.

[5] Z. Wagner *et al.*, “Women and children living in areas of armed conflict in africa: a geospatial analysis of mortality and orphanhood,” *The Lancet Global Health*, vol. 7, pp. e1622–e1631, 2019.

[6] I. Timæus and M. Jasseh, “Adult mortality in Sub-Saharan Africa: evidence from demographic and health survey,” *Demography*, vol. 41, 4, pp. 757–772, 2004.

[7] S. Helleringer *et al.*, “Improving the quality of adult mortality data collected in demographic surveys: Validation study of a new siblings’ survival questionnaire in Niakhar, Senegal,” *PLoS Med*, vol. 11, p. e1001652, 2014.

[8] V. Adjiwanou *et al.*, “Measuring excess mortality during the COVID-19 pandemic in low- and lower-middle income countries: the need for mobile phone surveys,” preprint 2020.

[9] I. Timæus, B. Zaba, and M. Ali, “Estimation of adult mortality from data on adult siblings,” in *Brass tacks: essays in medical demography* (B. Zaba and J. Blacker, eds.), pp. 43–66, Athlone Press, London, 2001.

[10] K. Wachter, D. Blackwell, and E. Hammel, “Testing the validity of kinship microsimulation,” *Journal of Mathematical and Computer Modeling*, vol. 26, pp. 89–104, 1997.

[11] E. Zagheni, “The Impact of the HIV/AIDS Epidemic on Kinship Resources for Orphans in Zimbabwe,” *Population and Development Review*, vol. 37, no. 4, pp. 761–783, 2011.

[12] I. Timæus, “Estimation of adult mortality from paternal orphanhood: a reassessment and a new approach,” *Population Bulletin of the United Nations*, vol. 33, pp. 47–63, 1992.

[13] H. Booth, “Transforming Gompertz’s Function for Fertility Analysis: The Development of a Standard for the Relational Gompertz Function,” *Population Studies*, vol. 38(3), pp. 495–506, 1984.

[14] W. Brass, “Perspectives in population prediction: illustrated by the statistics of England and Wales,” *Journal of the Royal Statistical Society*, vol. 137, pp. 532–583, 1974.

[15] K. Hill and J. Trussell, “Further developments in indirect mortality estimation,” *Population Studies*, vol. 31(2), pp. 313–334, 1977.

Acknowledgments

This study is supported by a grant from the Bill & Melinda Gates Foundation for the RAMMPS project (INV023211).