A method for estimating sub-national mortality rates by using incomplete civil registration to correct census data on recent deaths

Rob E. Dorrington[1] and Ian M. Timæus[2,1]

[1]CARE, University of Cape Town, South Africa, [2] Department of Population Health, London School of Hygiene & Tropical Medicine, UK

Abstract

It is notoriously difficult to estimate life tables which represent sub-national regions in developing countries. Death registration is usually incomplete and Death Distribution methods for estimating the completeness of the statistics often perform poorly at a sub-national level. On the other hand, census questions about recent deaths in the household are prone to unpredictable age-specific patterns of under- and over-reporting. This paper proposes a method of combining the two sources of data that makes use of the respective strengths of each to produce reliable estimates of sub-national mortality. It demonstrates the method by applying it to the estimation of provincial mortality in South Africa in 2001, 2006 and 2011 combining data on registered deaths, with those from two censuses and a large community survey.

Background

No accepted method exists for producing reliable estimates of mortality for sub-national/regional populations when vital registration is less than complete and/or does not record region of residence prior to death accurately. Because inter-regional migration can often be both substantial and difficult to estimate accurately, demographic Death Distribution methods (DDMs) for estimating the completeness of reporting of deaths tend to perform poorly at a sub-national level (Dorrington 2013a, b). In addition to this, people may die in regions other than where they are permanently resident (which is not often captured accurately by vital registration systems in developing countries).

Questions asking households to report deaths in the past year have been reintroduced into censuses in the 2010 round. While these questions ensure a high degree of correspondence between the deaths reported in a region and the regional population exposed to the risk of dying, the answers to these questions often produce unreliable estimates of mortality rates for other reasons. These include reference-period errors, whereby the deaths reported refer to either a shorter or longer period before the census than was intended, and both under-reporting of deaths (when households become extinct or disintegrate on the death of the person) and over-reporting (when deaths are reported by more than one household). As it seems likely that the deaths of elderly household heads are most likely to go unreported, and the deaths of young adults who are involved in labour migration are most likely to be reported by more than one household, the completeness of reporting may vary markedly by age.

Aims, data and methods

This research demonstrates that, by making use of civil registration and census data in combination, one is able to overcome the weaknesses of each, by estimating mortality in the provinces of South Africa in each of the years immediately prior to the 2001 census, the 2007 Community Survey (which administered a census questionnaire to a sample of a little over 2% of the population) and the 2011 census.

DDMs were applied to registered deaths and the census/survey populations by self-identified ethnic population group (Black African, Coloured, Indian and White) to estimate the completeness of reporting of mortality of those aged 5 or more at a national level by ethnic population group. The key assumption involved in this is that the completeness of death registration is the same at all ages within each population group. Completeness of vital registration of deaths under age 5 was estimated in a separate

exercise from that for population as a whole, that is, all ethnic groups combined (Bradshaw, Dorrington and Laubscher 2012), on the assumption that all under-reporting of deaths under age 5 occurred in the Black African population group[1].

These estimates were then used to produce national estimates of the expected number of deaths by sex and age group (0, 1-4, 5-9, …, 85+) for each ethnic population group.

Comparison of the number of deaths reported by households with the expected number provides age-sex-specific correction factors for the number of deaths reported for each population group in the censuses/survey. These correction factors were then applied to the number of deaths reported by households in each of the provinces on the assumption that the factors for each population group are the same for all provinces. The adjusted deaths for the four population groups were then combined and used together with census population numbers to estimate mortality rates in each province.

The reasonableness of the ranking of adult mortality and the reasonableness of the estimates at the older ages were checked by comparison with estimates of mortality produced using data on orphanhood.

Results and anticipated conclusions

Illustrative results are presented using the 2001 census data. The exercise has been repeated for 2006 using data from the 2007 Community Survey and for 2011 using data from the 10% public use micro-sample (PUMS) in combination with the cause-of-death vital registration data for 2006 and 2011, but only results for 2001 are presented in this abstract. (While the sample size of the Community Survey means that the results are more 'noisy', once smoothed, they produce mortality rates that are consistent with both those from the 2001 census as well as expectation given what is known about the level of the prevalence of HIV in each province).

Figure 1 shows the adjustment factors used to correct the numbers of deaths reported by households in the 2001 census, for each of the ethnic population groups in South Africa. While the overall level varies from one census to another, two distinct patterns are consistent through time. The first is the very noticeable under-reporting of deaths in the White population at the older ages. This reflects the high proportion of the population in this group who live (and die) either alone or in old-age homes. The second, less noticeable, phenomenon is the increasing under-reporting with age of the African females, which is probably the result, given the high proportion of female-headed households at these ages, of deaths going unreported due to the disintegration of the household on the death of these women.

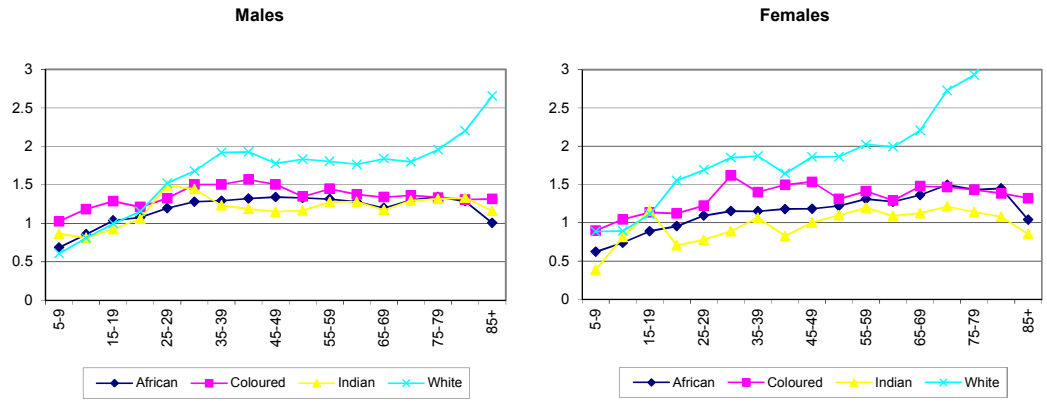

Figure 1 Factors needed to correct the number of deaths reported by households to produce the correct number of deaths by ethnic population group

[1] In turn based on the assumption that virtually all under-reporting occurs in the non-urban population, which is predominantly Black African.

Figure 2 presents mortality rates produced by the application of these factors to the census and survey reports. The rates exhibit the high mortality of young adults that is the distinctive 'signature' of HIV/AIDS mortality and rank the provinces as expected on the basis of existing knowledge of the HIV epidemic.

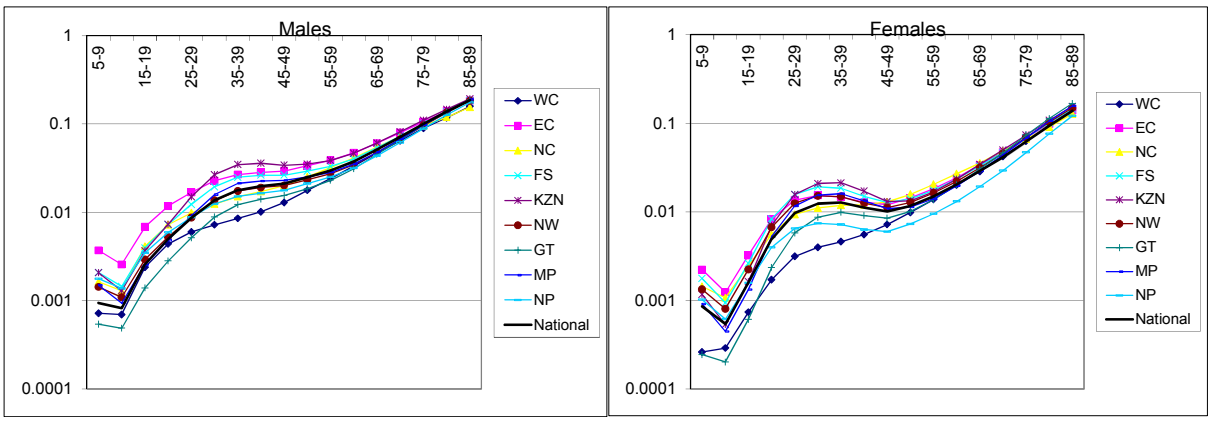

Figure 2 Male and female mortality rates by province for South Africa, 2001

Figure 3 compares the estimate of ${}_{15}q_{50}$ calculated from the above rates to that produced using the orphanhood data. Although a few of the points are quite far from the 45-degree line, this is to be expected, since the orphanhood estimates reflect mortality over a number of years before the census and the mortality of parents of people who live in a particular province is not exactly the same as that of people of the same age who live in the province (since some parents are likely to live in a province other than that of their children). Nevertheless, the two series are highly correlated and both the spread of the points based on orphanhood data and the national estimates are broadly consistent with the mortality rates produced using the deaths reported by households corrected for under or over-reporting.

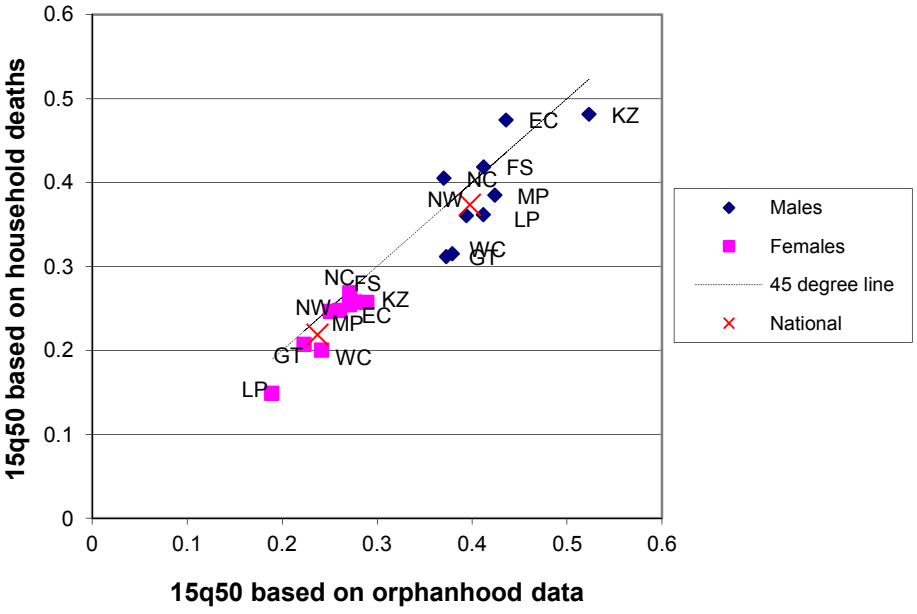

Figure 3 Comparison of estimates of 15q50 with those from the Orphanhood method

The conclusion is that the new method produces plausible results at the provincial level when applied to data from the two censuses and the Community Survey. In addition, the method provides useful evidence about the extent over/under reporting of deaths by households in censuses and similar inquiries.

Sub-national results could prove useful in improving various attempts to produce systems of model life tables. These often have to resort to using output from epidemiological projection models in the absence

of observed data (Sharrow, Clark and Raftery 2013; Wang, Dwyer-Lindgren, Lofgren *et al.* 2012). Our sub-national estimates, however, represent series of life tables that reflect not only differing levels of prevalence of HIV/AIDS, but also changes in prevalence over time in the same population.

Unfortunately, while the method appears to succeed at producing estimates at the level of province, it seems to work less well for populations below provincial level. This may be due to a combination of factors, including smaller populations being more subject to random fluctuations, less clear boundaries, and a greater likelihood of mismatch of place of death and place of residence.

Bradshaw, D., Dorrington, R. E. and Laubscher, R. 2012. *Rapid Mortality Surveillance Report 2012*. Cape Town: South African Medical Research Council. http://www.mrc.ac.za/bod/RapidMortality2011.pdf

Dorrington, R. E. 2013a. "The Generalized Growth Balance method." in Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K. H., Timæus, I. M. and Zaba, B. (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population. Available: http://demographicestimation.iussp.org/content/general-growth-balance

Dorrington, R. E. 2013b. "The Synthetic Extinct Generations method." in Moultrie, T. A., Dorrington, R. E., Hill, A. G., Hill, K. H., Timæus, I. M. and Zaba, B. (eds). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population. Available: http://demographicestimation.iussp.org/content/synthetic-extinct-generations

Sharrow, D. J., Clark, S. J. and Raftery, A. E. 2013. "Modeling Age-Specific Mortality for Countries with Generalized HIV Epidemics." Paper presented at the Annual meeting of the Population Association of America. New Orleans, USA, 11-13 April. Available:

Wang, H., Dwyer-Lindgren, L., Lofgren, K. T., Rajaratnam, J. K., Marcus, J. R., Levin-Rector, A., Levitz, C. E., Lopez, A. D. and Murray, C. J. L. 2012. "Age-specifi c and sex-specifi c mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study 2010", *Lancet* **380**:2071-2094.