Extended Abstract Submitted to the Annual Meeting of the Population Association of America - 2012

Estimating level and structure of mortality in small areas: counterfactual approaches

Bernardo Lanza Queiroz

Everton Emanuel Campos de Lima

Flávio Henrique Miranda de Araújo Freire

Diana Oya Sawyer

Josivan Ribeiro Justino

Background

In developing countries, mortality estimates and knowledge of levels and trends of mortality are limited by the quality of data. The most common problems faced in these countries are incomplete coverage of vital registration systems and errors in age declaration for both population and death counts. In recent years, collection of data for death counts has improved, but there are still limitations for studying mortality in several parts of the world. The problem is more complex when it is studying small areas and sub-national populations. Generally, the accurate estimation of rate schedules for small areas has become more important as demographers have gained greater access to geocoded data. However, even with very large samples and censuses, small areas often have small risk populations that produce unstable estimates. Furthermore, the mortality data for small areas are affected by the same problems mentioned before and by regular fluctuations (small numbers) in the region. That is to say, with common estimation techniques, the small-populated areas often produce extreme estimates, dominated by sampling noise, that may have little relationship to underlying local risks (Bernadinelli and Montomoli 1992).

The most common alternative for studying mortality in settings where data is limited is to produce estimates using demographic methods, such as Brass indirect techniques and/or death distribution methods. However, it is argued that these methods do not work properly in small areas. Thus, public health administrations are faced with limited information to allocate resources and it is also difficult to study the progress of public policy interventions at the sub-national levels, limiting the action of government agencies in improving the quality of life of these sub-populations. In this context, we aim to study the quality of mortality data for small areas in Brazil and estimate the level and structure of mortality using counterfactual approaches. The main idea is to compare different approaches of mortality estimation, in order to get a sort of confidence interval of mortality schedules for least-populated areas.

Data and Methods

We make extensive use of Datasus, the brazilian vital statistics database. Datasus has information on death counts by age and sex (and causes) for all municipalities in Brazil since 1979. We use data from the last decade to perform this study by both death and population counts.

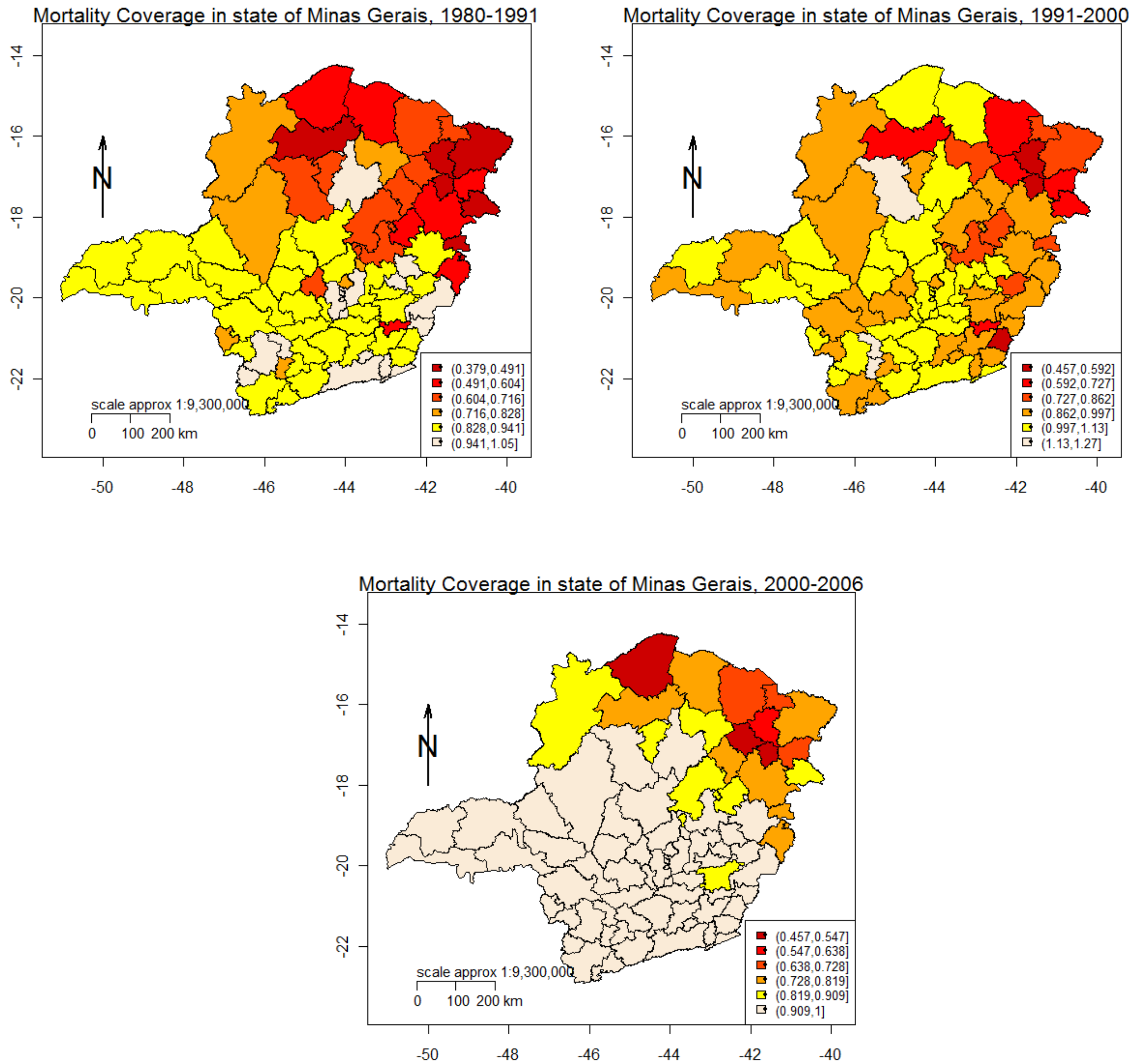
In this paper, we proposed a combination of standard demographic methods (Death Distribution Methods) with empirical Bayesian (EM) statistics and Expectation Maximization (EM) algorithm to produce estimates of levels and structure of mortality for small areas in Brazil. First, we begin by producing mortality estimates using traditional demographic methods and providing a descriptive analysis of the results and its limitations. The next step, we then produce separate estimates using empirical Bayesian and demographic standardization methods and, finally, we compare the results and applicability of both methods. We then move on to propose a combination of demographic methods and statistical methods to study mortality levels and structure.

In this paper, we focus on small areas for 3 states in the country: Minas Gerais (MG), Rio Grande do Norte (RN) e Rio Grande do Sul (RS). We chose these three regions to be able to test our estimates under different data quality situations: RS is considered to have complete coverage, thus we will test the method without the need to adjust for under-registration of deaths counts; MG and RN combine small areas that are considered to have good data quality with others with worse quality. In this way, we expect to have a big picture of how our method is working while it is applied to different demographic contexts.

Preliminary results and findings

In attempting to compare mortality estimates in least-populated areas, first we analyze the death counts under-registration in small areas of the states of RN and MG. Taking into account the next figure 1, we can see that the death counts estimates for MG – project by traditional demographic methods – show an expressive gain in mortality coverage, presenting considerable reduction and concentration of death counts under-registration at the poorest and less developed state areas of Northern and Northeastern.

Figure 1: Mortality coverage evolution – State of Minas Gerais, 1980 to 2006



Source: Lima and Queiroz, 2010.

For future mortality estimates, we presume that will be necessary to correct the death counts under-registration for the sub-populations from Northern parts of this state.

Regarding the state of RN, in order to evaluate the death counts we employed two Bayes statistic methods, respectively the Empirical Bayes[1] and Expectation Maximization. This time, the mortality estimates consider the relationship between observed and expected number of deaths, estimated according to a common mortality risk applied to all small areas. This general risk is projected according to the death count information of a larger geographical area.

In the Census year 2000, for the state of RN as a whole, we estimate 11% to 13% of mortality underreporting[2], depending on considered method. The estimates show different results as we consider sub-state populations. The Eastern parts of the state – where the capital and other major cities are located – are characterized by highest urbanization rates and lowest levels of mortality under-registration. In these areas, the levels of deaths coverage are somewhat above 90%. However, it is verified a large the spatial heterogeneity with respect the coverage of death counts across small areas. That is to say, in the Eastern parts of the state we can find small locations with 100% as well as barely 31% of recorded deaths.

These are preliminary results, subsequently, we intend to evaluate the mortality counts for the state of RS, although, this time using the two estimation methods, i. e. traditional demographic methods and Bayes statistics. Once we estimate the mortality under-reporting for the three states, the next step is to estimate life tables for the sub-population in each state.

References

Ahmed, S. E Hill, K. (2011). Maternal mortality estimation at the subnational level: a model-based method with an application to Bangladesh. **Bulletin of the World Health Organization**, 89:12-21.

---

[1] This Empirical Bayes (EB) technique is seen as a shrinkage method that reduces the mortality fluctuations of small risk populations based on mortality information of a large area.

[2] We assume that the obtained differences between observed and expected could be interpreted as underreported mortality.

Assunção, Renato M., Schmertmann, Carl P., Potter, Joseph E. Cavenaghi, Suzana M. (2005). Empirical Bayes Estimation of Demographic Schedules for Small Areas. **Demography**, Vol. 42, No. 3, (Aug., 2005), pp. 537-558

Banister, J.; Hill, K. Mortality in China 1964-2000. **Population Studies**, 2004, v. 58, n. 1, p.55-75

Bennett N.G. and Horiuchi S. (1981) Estimating the Completeness of Death Registration in a Closed Population. **Population Index** 47(2):207-21.

Bernadinelli, L. and C. Montomoli. 1992. "Empirical Bayes Versus Fully Bayesian Analysis of Geographical Variation in Disease Risk." Statistics in Medicine 11:983-1007.

Bhat, Mari. General growth balance method: a reformulation for populations open to migration. **Population Studies**, 56 (2002), 23-34

Brass W. (1964) Uses of census or survey data for the estimation of vital rates" (E/CN.14/CAS.4/V57), Paper prepared for the African Seminar on Vital Statistics, Addis Abbaba . 14-19 December.

Brass, William. **Methods for Estimating Fertility and Mortality from Limited and Defective Data**. Chapel Hill, N.C: International Programme of Laboratories for Population Statistics, North Carolina Population Center, 1975.

Dempster, A. P.; Laird, N. M.; Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society B**, v. 39, n. 1, p. 1–38, 1977.

Freire, F. H. M. A. . Projeção populacional para pequenas áreas pelo método das componentes demográficas usando estimadores bayesianos espaciais. Belo Horizonte: CEDEPLAR, 2001 (Tese de Doutorado).

Hill K, (1987). “Estimating census and death registration completeness.” ***A*sian and Pacific Census Forum**, 1(3): 8-13,23-24.

Hill, K. (2003) Metodos para estimar la mortalidad adulta en los paises en desarollo: una revision comparativa. **Notas de Poblacion**, no. 76 pps. 81-111.

Hill, K.,, You, D. e Choi, Y. (2009) Death Distribution Methods for Estimating Adult Mortality: sensitivity analysis with simulated data errors **Demographic Research,** vol 21, 2009 .

Hill, Kenneth. *Methods for Measuring Adult Mortality in Developing Countries: A Comparative Review.* The Global Burden of Disease in Aging Populations – Research Paper No. 01.13, 2000.

Hill, K.; Choi Y.; Timaeus, I. Unconventional approaches to mortality estimation. **Demographic Research**, v.13, 2005, p.281-300.

Luy, M. (2010). A Classification of the Nature of Mortality Data Underlying the Estimates for the 2004 and 2006 United Nations' World Population Prospects. **Comparative Population Studies,** vol. 35, no. 2, 2010.

Timaeus, I. M. Measurement of adult mortality in less developed countries: A comparative review. **Population Index***,* v.57, n.4, p.552-568, 1991

Zlotnik H and Hill KH. (1981). "The Use of Hypothetical Cohorts in Estimating Demographic Parameters Under Conditions of Changing Fertility and Mortality." **Demography** 18(1):103-22.