MORTALITY ESTIMATES AND CONSTRUCTION OF LIFE TABLES FOR SMALL AREAS IN BRAZIL, 2010

Flávio H.M.A. Freire
Universidade Federal do Rio Grande do Norte
fhfreire@ccet.ufrn.br

Bernardo L. Queiroz
Universidade Federal de Minas Gerais
lanza@cedeplar.ufmg.br

Everton E.C. Lima
Universidade Estadual de Campinas
everton@nepo.unicamp.br

Marcos R. Gonzaga
Universidade Federal do Rio Grande do Norte
mrcs.roberto@gmail.com

Felipe H. de Souza
Universidade Federal do Rio Grande do Norte
felipehenrique87@yahoo.com.br

Acknowledgments

Financial support from CNPq and Fapesp is greatfully acknowledge.

Structured Abstract

BACKGROUND
The study of mortality level and trends in developed countries is limited by the quality of the vital registration system. 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.

OBJECTIVE
We aim to study the quality of mortality data for small areas in Brazil and estimate levels and structure of mortality combining demographic methods with empirical Bayesian methods. We also produce estimates of life tables, by sex, for all municipalities in Brazil in 2010.

METHODS
We first use Death Distribution Methods (DDM) to analyze the quality of mortality data and estimate adult mortality levels for microregions in Brazil. We then combine DDM with empirical bayes approach to produce mortality estimates for small areas (municipalities) in Brazil. The two step procedure, allows us to correct the underegistration of death counts and minimize random fluctuations that might occur when we calculate mortality levels in small areas. The procedure to estimate mortality under age 15 uses relational models to correct the underegistration of death counts and to produce complete life tables for males and females.

RESULTS
The results show that completeness of death counts in the country is almost 100% in 2010, compared to 85% in the 1990s, but there is still a lot of variation across regions. More developed regions have complete registration whereas less developed areas in the North and Northeast have low registration. We also find a clear spatial pattern of mortality across regions in Brazil. For males, preliminary results indicates six major groups of mortality patterns. For females, we estimate seven different mortality patterns groups.

CONCLUSIONS
We find that the quality of mortality data in Brazil and regions is improving over time, a large part of the country shows almost complete coverage of death counts. We also estimate mortality levels for all municipalities in Brazil in 2010, over 5000 estimates, and observed the existence of clear regional spatial patterns of mortality. We are working with different models to check the robustness of our results and comparing our results with estimates produced by UNDP in 2010. We believe our estimates are more robust and accurate since we make use of information from all age groups, whereas UNDP estimates life expectancy based only on infant and child mortality and Coale-Demeny model life tables.

Keywords: small areas, mortality estimation, demographic techniques, empirical bayessian statistics

Introduction

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. In Brazil, mortality estimates and the knowledge of levels and trends of mortality are limited by the quality of data (França, et.al, 2012; 2008; Paes, 2005; 2007; Gomes e Turra, 2009). The most common problems faced are incomplete coverage of vital registration systems, errors in age declaration for both population and death counts, and lack of information on causes of deaths. These limitations are even more striking in small areas of the country, such as municipalities or counties (Cavalini, et.al, 2007; França, et.al, 2008; Paes, 2007). According to international standards, Brazil is characterized by regular levels of coverage of death counts, but with large variation across regions (Paes, 2005; Luy 2010, Setel et.al, 2007, PAHO, 2010; The PLOS, 2010).

The inability to produce proper estimates of mortality, especially in small areas, harms the development of public health policies and the understanding of the health transition in the country. On one hand, a lot is known about variations of infant and child mortality in Brazil (Souza, Hill and Dal Poz, 2010; Castro and Simões, 2009), but, on the other hand, very little is known about spatiotemporal trends in adult mortality in Brazil. We argue that producing proper estimates of adult mortality for small areas in Brazil is very relevant because recent and future changes in life expectancy are probably going to be explained by variations for adults and the elderly, since there is a clear trend in convergence in the levels of infant and child mortality (Souza, Hill and Dal Poz, 2010).

In recent years, the accurate estimation of rate schedules in small areas has become more important as demographers and health experts have gained greater access to geocoded data. However, even with very large samples and censuses data, 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 traditional demographic techniques, the small-populated areas often produce extreme estimates, dominated by sampling noise that may have little relationship to underlying local mortality risks (Bernadinelli and Montomoli 1992).

Although, the search for better estimates of rates in small areas gained ground in the research agenda of many demographers even with large samples and censuses, vital rates estimates in small areas are still very limited and incipient. This often happens due to the problem of few events recorded in the denominator and/or numerator of the measures of interest. This instability is even worse when sub-national groups are disaggregated by age and sex (Assunção et al. 2005). Nevertheless, Bernadinelli and Montomoli (1992) argue that, in small populations, the estimated rates generally have extreme values, often dominated by sampling noise which less reflect the true risks. Assunção et al. (2005) also argues that for a large number of small areas, one can observe a large variability in the estimated rates that do not reflect the true level of heterogeneity of the geographic location. Therefore, estimates of vital rates in small areas present a great challenge for demographers, but several authors argue that a variety of statistical methods exist to adequately address the volatility of these estimates (for example Ferguson, 2004). In studies estimating fertility rates in Brazil, Assunção et al. (2005), using the empirical bayesian estimaton, showed how this methodology was effective in the case of Brazilian municipalities. According to them, empirical Bayesian estimations are in many cases better than traditional demographic estimates, especially when the studied phenomenon is characterized by a strong spatial and age patterns (Assunção et al., 2005; Schmertmann, 2013, among others).

In this context, incomplete data and sub-notifications, knowing the level and pattern of mortality and trying to understand the historical trend becomes a difficult exercise insofar as there is a difference in coverage and quality of the data over time and regions. Moreover, at the same time increases the demand for socio-economic and demographic information in disaggregated geographical level. Therefore, the need for reliable information of death in sub-national levels of disaggregation is increasing. One needs to know the pattern and level of mortality in Brazilian municipalities. Does within a state does not have distinct patterns of mortality? Mortality in middle region of a state, is not much bigger than the other? Questions like these need to be answered.

The Atlas of Human Development, introduces municipal mortality indicators: infant mortality; Child mortality (children up to 5 years); life expectancy at birth;

probability of survival to age 40 and the probability of survival to age 60. However, the methodology used for the construction of these indicators at the municipal level, use the method of infant and child mortality developed by Brass (1968) combined to Coale-Demeny model life tables. We argue that those methods do not work properly in small areas and in situation of rapid mortality decline. Thus, public health administrations are faced with limited information to allocate resources and it is 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 paper, the focus is producing more accurate estimates of the death counts underregistration and to estimate adult and child mortality for every Brazilian municipality, for each sex and five year age group. The central methodological idea of this paper is to combine traditional demographic methods (Death Distribution Methods) with empirical Bayes statistics. That is, we first correct the undercounts of death counts and to estimate mortality combining statistical approach to demographic methods in the small areas estimates. With this, we correct the underreporting of deaths and minimized random fluctuations that occur when estimates are any indicator in small areas, with small numbers, with low population and low number of contingent events.

Methodology

We first use Death Distribution Methods (DDM) to analyze the quality of mortality data and estimate adult mortality levels for microregions in Brazil. We then combine DDM with empirical bayesian approach to produce mortality estimates for small areas (municipalities) in Brazil. The two-step procedure, allows us to correct the underegistration of death counts and minimize random fluctuations that might occur when we calculate mortality levels in small areas. The procedure to estimate mortality under age 15 uses relational models to correct the underegistration of death counts and to produce complete life tables for males and females. In sum, 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.

We make extensive use of the Ministry of Health database, DATASUS (http://www.datasus.gov.br). The data are collected at municipality level, but we aslo aggregated them in comparable small areas, using the National Statisitcs Office (IBGE) definition of comparable mesoregions. Population by age and sex come from the Population Censuses (2000), collected by the Brazilian National Statistics Office (IBGE).

Adult Mortality Estimation: mesoregions

Correction of underreporting of deaths to estimated future adult mortality, developed in this work part of mesoregion first, and then fix and adjust municipal deaths and estimate adult mortality of Brazilian municipalities. First an indirect standardization of mortality rates by age was taken, according to sex, the mesoregion (smaller area), using as the standard structure of mortality by age and sex of the largest reference area, in this case the federal state (UF) which one belongs mesoregion. Thus, it eliminates the first inconsistencies that occur in small numbers, which makes the mortality rates by age and sex of some very unstable mesoregion. To those expected death counts in each mesoregion, after indirect standardization, we applied the adjustment factor for the under-registration of death counts using the death distribution methods.

Several methods based on equations of population dynamics have been developed to evaluate the coverage of reported deaths relative to population. The death distribution methods (DDM) are commonly used to estimate adult mortality in a non-stable population (Timeaus, 1991; Hill et al, 2005). There are four major approaches: the General Growth Balance (GGB) Methods (Hill, 1987), the Synthetic Extinct Generation (SEG) method (Benneth & Horiuchi, 1981), the Adjusted Synthetic Extinct Generation (SEG-adj) method (Hill, You & Choi, 2009), and the Synthetic Extinct Generation plus delta (Dorrington, 2011). The death distribution methods make several strong assumptions: that the population is closed to migration that the completeness of recording of deaths is constant by age, that the completeness of recording of population is constant by age, and that ages of the living and the dead are reported without error.

In this paper, we estimate coverage of deaths using the adjusted Synthetic Extinct Generations approach (SEG-adj) as proposed by Hill, You and Choi (2009). Hill, You and Choi (2009) suggest a combination of the methods of Hill (1987) and Bennett and Horiuchi (1981), which can be more robust than the application of two methods separately. The adjusted method is the application of general equilibrium (Hill,

1987) equation to obtain estimates of the change in census coverage (k1 / k2) and use these estimates to fit a demographic census (enumeration of the population) and then apply the method Bennett and Horiuchi (SEG) using the set for the coverage of mortality data population.

Small Area estimation: adult mortality

We cannot apply SEG-adj to small areas in Brazil because several assumptions of the method are not met, especially closed population. In order to overcome this limitation, we apply an indirect standardization by multiplying the mortality rate of the larger area (mesoregio), after the adjustment described above by the population of the smaller area (municipality). By doing this, we get an estimate of the expected deaths for each city-level i (ESP i). Therefore, besides the undercount, correction already made in mesoregion and incorporated into the municipal level by indirect standardization, it is advisable to make another adjustment now concerning themselves with the greater robustness of the estimated number of deaths by age group and sex.

Then, to minimize the variability in the estimate of the deaths of the municipalities we use empirical Bayesian estimator in estimates of mortality for each five-year age group and by sex. This estimator works as a sort of weighted average of correction for underreporting.

Small area estimation: under 15 years old mortality

The final step is to produce mortality estimates for the population under 15 years old. We used regression model approach to estimate the probabilities of death from 0 to 14. In the model, we used different estimates of adult mortality selected by statistical approach and controlling for multicollinearity. We run a different model for each 5-year age group (0-4, 5-9 and 10-14).

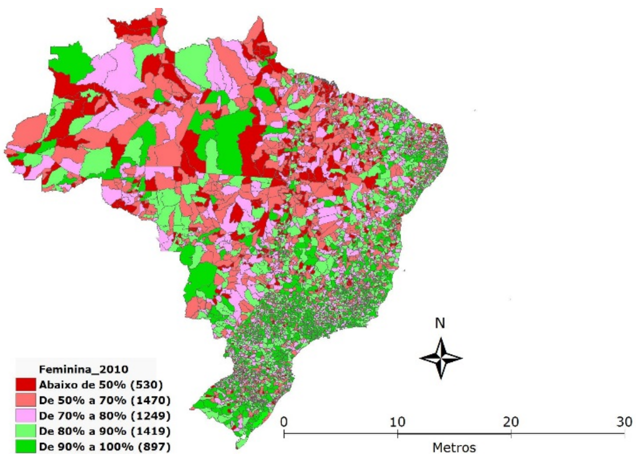
In order to estimate the coefficients we used different mortality patterns based on different countries with data available in the Human Mortality Database (www.mortality.org). In our case, for each Brazilian region we found a pattern of mortality in the HMD that resembled the experience of that particular region. Based on the HMD country data we estimated the regression coefficient for each different region. We then used the estimates of adult mortality, obtained before, as an explanatory variable to produce estimates of mortality under 15 for each municipality.

Preliminary Results

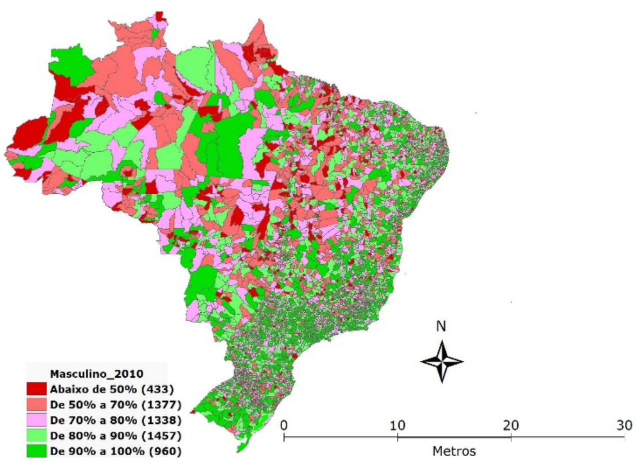
The results show that completeness of death counts in the country is almost 100% in 2010, compared to 85% in the 1990s, but there is still a lot of variation across regions. More developed regions have complete registration whereas less developed areas in the North and Northeast have low registration. We also find a clear spatial pattern of adult mortality across regions in Brazil. For males, preliminary results indicates six major groups of mortality patterns. For females, we estimate seven different mortality patterns groups. The patterns were built using cluster analysis based on the estimates of 45q15 for each municipality.

Map 1 – Death counts coverage, Brazil, males and females, 2010.

Panel A – Females

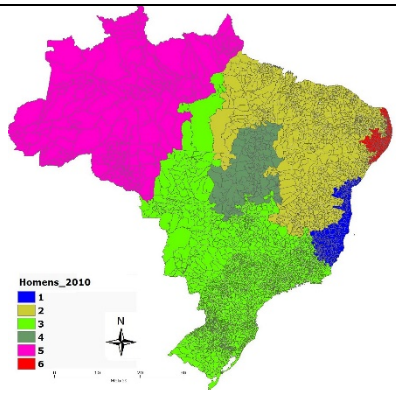
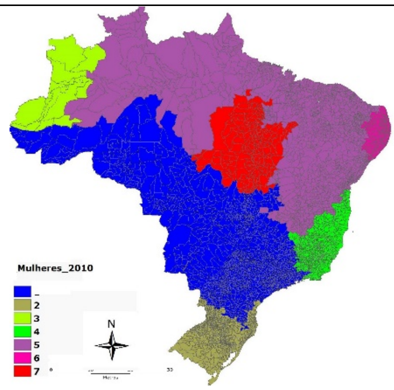


Panel B – Males



Source: Datasus and IBGE, 2010.

Map 2 – Small area mortality patterns, Brazil, males and females, 2010.

Males	Females
	

Source: Datasus and IBGE, 2010.

Discussion

We find that the quality of mortality data in Brazil and regions is improving over time, a large part of the country shows almost complete coverage of death counts. We also estimate mortality levels for all municipalities in Brazil in 2010, over 5000 estimates, and observed the existence of clear regional spatial patterns of mortality. We are working with different models to check the robustness of our results and comparing our results with estimates produced by UNDP in 2010. We believe our estimates are more robust and accurate since we make use of information from all age groups, whereas UNDP estimates life expectancy based only on infant and child mortality and Coale-Demeny model life tables.

The results presented are preliminary, we are testing different models and approach to estimate mortality levels for those under 15 years of age. We are also working to estimate life tables using the two-dimensional mortality model, proposed by Wilmoth and colleagues, to obtain estimates for each municipality in Brazil. We are also comparing estimates produced by UNDP to our estimates and test the robustness of the method proposed here.

References

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.

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.

Castro, M.C and Simões, C. (2009). Spatio-Temporal Trends of Infant Mortality in Brazil. Paper presented at the XXVI IUSSP International Population Conference.

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.

Dorrington R.E. General Growth Balance. In Moultrie TA, RE Dorrington, AG Hill, KH Hill, IM Timæus and B Zaba (eds), Tools for Demographic Estimation. http://demographicestimation.iussp.org/content/general-growth-balance, 2014

Dorrington R.E. Synthetic extinct generations. In Moultrie TA, RE Dorrington, AG Hill, KH Hill, IM Timæus and B Zaba (eds), Tools for Demographic Estimation. http://demographicestimation.iussp.org/content/synthetic-extinct-generations, 2014

Ferguson, B.; Reniers, G.; Araya, T.; Jones, J. H.; Sanders, E. Empirical Bayes Estimation of Small Area Adult Mortality Risk in Addis Ababa, Ethiopia. In: Annual Meeting of the Population of Association of America – 2004. Boston, Massachusetts, April 1-3, 2004.

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).

Gelman, A. "Bayesian data analysis". London; New York: Chapman & Hall,1995.

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

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

Kintner, H. J. 1988. "Determinants of Temporal and Areal Variation in Infant Mortality in Germany, 1871-1933". Demography 25 (4): 597–609.

Klüsener, Sebastian, and Rembrandt D. Scholz. "Regional hot spots of exceptional longevity in Germany." Vienna Yearbook of Population Research 11 (2013): 137-163.

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.

Marshall, R. M. Mapping disease and mortality rates using Empirical Bayes Estimators. Journal of the Royal Statistical Society, Series C: Applied Statistics, v. 40, n. 2, p. 283-294, 1991.

Ministério da Saúde. Sistema de Informações sobre Mortalidade (SIM) de 1979 a 2010.

Murray CJL, Rajaratnam JK, Marcus J, Laakso T, Lopez AD. What can we conclude from death registration? Improved methods for evaluating completeness. PLoS Medicine. 2010 April 13; 7(4)

Paes, N. A. Avaliação da cobertura dos registros de óbitos dos Estados brasileiros em 2000. Revista de Saúde Pública, USP, v. 39 n.6, p. 882-90, 2005.

Paes, N. A.; ALBUQUERQUE, M. E. E. Avaliação da qualidade dos dados populacionais e cobertura dos registros de óbitos para as regiões brasileiras. Revista de Saúde Pública, USP, v. 33, n. 1, p. 33-43, 1999.

Pollard, A. H. Random Mortality Fluctuations and the Binomial Hypothesis, Journal of the Institute of Actuaries (JIA), 96 (1970) 251-264.

Potter, Joseph E., Carl P. Schmertmann, and Suzana M. Cavenaghi. "Fertility and development: evidence from Brazil." Demography 39, no. 4 (2002): 739-761.

Potter, Joseph E., Carl P. Schmertmann, Renato M. Assunção, and Suzana M. Cavenaghi. "Mapping the timing, pace, and scale of the fertility transition in Brazil." Population and development review 36, no. 2 (2010): 283-307.

Schmertmann, Carl P., Joseph E. Potter, and Suzana M. Cavenaghi. "Exploratory analysis of spatial patterns in Brazil's fertility transition." Population Research and Policy Review 27, no. 1 (2008): 1-15.

Schmertmann, Carl P., Suzana M. Cavenaghi, Renato M. Assunção, and Joseph E. Potter. "Bayes plus Brass: Estimating total fertility for many small areas from sparse census data." Population Studies 67, no. 3 (2013): 255-273.

Sparks, Patrice Johnelle, and Corey S. Sparks. "An application of spatially autoregressive models to the study of US county mortality rates." Population, Space and Place 16, no. 6 (2010): 465-481.

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

Wilmoth, J.; Zureick, S.; Canudas-Romo, V.; Inoue, M.; Sawyer, C. (2012). A flexible two-dimensional mortality model for use in indirect estimation. Population Studies, 66(1): 1-28.