Mortality Estimates for Small Areas in Argentina (2009-2011)

Nicolás Sacco;[*] *Iván Williams;*[†] *Bernardo L. Queiroz*[‡]

Abstract

The lack of mortality data available for developing nations have prevented them from fully study causes and consequences of trends and changes. To cover such issue many countries have relied on specialized surveys, that have proven ideal for national estimates and related variables. However, sub national mortality trends and its associated variables by today remain a big unknown. Most Latin American Countries are now capable of producing such data, that even with large under registration issues constitute a first step to produce this very relevant data that will allow the study of mortality related to important matters, including public health and policy related purposes. With several countries in the region approaching high levels of life expectancy at birth, comparative research on the region is a pressing demand for population scientists. The experience of data harmonization efforts such as the World Health Organization Mortality Database, Human Mortality Database, the Latin American Mortality databases, LAMDA and LAHMD and the Global Bilateral Migration database, to mention but a few of examples, has proven to be highly beneficial for comparative studies. In the realm of mortality, the Human Mortality Database constitute a remarkable example, yet, it only includes one Latin American country: Chile. More recently, an increasing interest for exploring sub-national variation in demographic outcomes started to play a major role in demographic studies and data harmonization efforts. Exploring sub-national heterogeneity in mortality is a promising research venue in Latin America given the large socioeconomic inequality that characterizes the region. In this context, in this article we estimate mortality levels for small areas in Argentina (selected provinces) during the period 2009-2011. For this, we use bayesian empirical estimation. The work plan will allow building an input both for the elaboration of public policies and for deepening the differentials of mortality in the selected geographical areas and to contribute for the upgrade of a harmonized and comprehensive data base of mortality indicators for different sub-national levels, the LAHMD.

Keywords: Mortality; Argentina; Small areas.

Introduction

Despite extensive work in developed countries, little is known about mortality differences in small areas over the Southern Cone. In particular, in Argentina, mortality estimates and the knowledge of their levels and trends are limited, as in most countries of the region, by the quality and availability of data. Most frequent problems are associated with incomplete

[*]Population Research Institute - Penn State University, nsacco@psu.edu

[†]Universidad Nacional de Luján, ivanwilliams1985@gmail.com

[‡]Cedeplar-UFMG, lanza@cedeplar.ufmg.br

coverage of the vital statistics system, errors in the declaration of age in the populations involved in the calculation of rates and the lack of information on causes of death.

Recent studies focused on Brazil (Lima, Queiroz, and Sawyer 2014; Lima and Queiroz 2014; Freire 2015; Queiroz et al. 2017) have applied different methodologies to estimate mortality in the sub-national horizon. There are reasons to suspect that what is observed in Brazil can also occur in Argentina, despite the fact that both countries went through very different epidemiological transition processes.

Considering that there are no precedents in the subject for the Argentine case, in this paper we estimate levels and structure of mortality in minor administrative areas in Argentina, during the period 2009-2011. To that end, we observed mortality changes in selected states: Catamarca, Neuquén, Córdoba, Salta, Interior Buenos Aires and the 24 counties in the Greater Buenos Aires.

Data on deaths and causes of death are essential to establish priorities for investment in public services, the application of economic planning policies and their monitoring, at the national level and, increasingly, at lower administrative levels (Setel et al. 2007). In Latin America and the Caribbean this type of information, at the local level, is scarce or non-existent -even taking into account ad-hoc surveys- and the demand for this data is growing, both as an input for the application of different development plans as for the allocation of resources.

Literature Review

Since 1914, the continuous rise in period life expectancy at birth (LEB), distinguished Argentina from the rest of Latin America. Due to early socio-economic development, high degree of urbanization and the expansion of formal education, mortality rates have been substantially reduced (Recchini de Lattes and Lattes 1975; Grushka 2014). After moderate at the beginning of the 20th century, a steady growth in LEB was distinguished and continued to increase until the 1960s. According to official data, this regular growing behavior, in both sexes, was interrupted around 1970 with an apparent setback of 1.8 years, compared to the previous decade. This phenomenon has not been adequately addressed yet. Müller and Accinelli (1980) attributed it to an arrival at a threshold in the mortality gains imposed by the socioeconomic conditions. This hypothesis suggested that health progress lost its independence from levels of economic development.

This interpretation was based on the idea that the main ceilings reached in LEB values were found in the regions with the highest socio-economic development, while in the rest of the areas, which are relatively less developed, evidence a rise in LEB. The 1969-1971 mortality table showed a mortality increase that affected especially males. Male LEB decrease with respect to that recorded for the period 1959-1961 in the regions of greater relative development, while it increase in less developed regions, where the level of LEB was lower in the previous period. Although moderate compared to previous years, the advances in the LEB that would occur in the subsequent period, despite the deterioration of the socioeconomic conditions of Argentina as a whole until 2003, would show the limits of such hypothesis. Since 1980, LEB begins again trend consistent with the decline in mortality. This period was

followed by a recovery of the upward trend, but in a slower way, reaching a LEB of 75.34 years (both sexes) in 2009.

Hypothesis

As a general framework, Argentina can be contextualized, starting from its modernization and successful integration with the world market at the end of the 19th century, as an anomalous country in demographics and health terms: an important recipient of immigrants with a broad coverage of the education system, and where low unemployment, low cost of food and the development of effective sanitation in the cities produced a good sanitary level, and consequently a low mortality rate, both infant and adult. However, this path -a global pioneer- slows down since the mid-1970s.

This article proposes to contrast the link, at a dis-aggregated level, between levels and structure of mortality at the sub-national level that the theoretical approximations pondered in different measure as linked, in the long term, with economic development. What role plays the dynamics of small area mortality in the configuration of mortality levels at the regional and total country levels? Is there a deceleration of the decline in mortality in Argentina due to the co-existence of socially heterogeneous groups, lagging behind in the process of epidemiological transition? Are differential patterns associated with different economic development models applied throughout different sociology-historical cycles? Main hypothesis: the levels of mortality and the rate of decline, at the provincial and total country levels, are dependent on the growing social inequality, a possible aspect to be glimpsed through the sub-national observation of mortality. Based on these questions and working hypotheses, in this paper we study the levels and structure of mortality at the sub-national level (departments) in relation to two specific dimensions: a) as an element in the long-term evolution of social and economic development and b) as a factor associated with different stages of the demographic and epidemiological transition.

Data & Methods

We used official data on death elaborated by the Health Secreatary (DEIS).

Although Argentina is usually classified as a country with good death statistics, it is known a unequal percentage of child deaths not registered by province (source DEIS). In this sense, since the data source is a record, considering the estimation of mortality rates by age, we could think that the variance would be null and the bias (both components of the mean squared error of an estimator) would be given by the pattern of cases omitted in each jurisdiction. The second component of the error will not be addressed in this paper, although there are increasing attempts to address the issue, depending on the available auxiliary information (Schmertmann and Gonzaga 2018; Queiroz et al. 2017) (nor the one that refers to the problems of census counting).

Regarding the first issue, it is not possible to assume is null because there are events with a small number of "experiments" (few exposed in our case), which present a greater variance

in their estimates, so it requires special treatment in order to reflect the risk of underlying mortality. To this end, we use the empirical Bayesian method, in order to improve the statistical efficiency of mortality rate estimators, decreasing small area variance (Efron and Morris 1972; Marshall 1991; Longford 1999; Assunção et al. 2005). The a prior distribution corresponds to the joint distribution of the mortality rates vector by age of the largest area. Then, through the observed in each small area, the Bayesian adjustment of the a posterior mortality distribution occurs. The characteristic of "empirical" is that the distributions of the parameters of the largest area are estimated from the data.

The definition of greater area must exploit the idea of internal similarity, to be able to suppose that the mortality of smaller areas are realizations of a greater stochastic process, supposedly important. The similarity in mortality patterns should be approached indirectly, which in our case, and because it is an incipient investigation, is supposed to belong to the same province, an issue that is not always correct, where the "distance" between jurisdictions is not it is measured by kilometers (Longford 1999).

Considering a five-year age group either in an area i, the distribution of deaths d a Poisson process is assumed, where $E(d^i_{x,4}/m^i_{x,4}) = N^i_{x,4}m^i_{x,4}$, being N those exposed and m the mortality rate. Being $\hat{m}^i_{x,4}$ the maximum likelihood estimator $D^i_{x,4}/N^i_{x,4}$ and given an a prior distribution of the parameter $m_{x,4}$ then his unconditioned expected value would be $E_m(E(\hat{m}^i_{x,4})/m^i_{x,4}) = E_m(m^i_{x,4}) = m_{x,4}$, and its unconditional variance $V_m(E(\hat{m}^i_{x,4}/m^i_{x,4})) + E_m(V(m^i_{x,4}/m^i_{x,4})) = V_m(\hat{m}^i_{x,4}) + E_m(\frac{m^i_{x,4}}{N^i_{x,4}}) = V_m(m^i_{x,4}) + \frac{m_{x,4}}{N^i_{x,4}}$.

The Bayesian linear estimator $\mathring{m}^i_{x,4}$ that minimizes the mean square error of $m^i_{x,4}$ and indexes that are linear functions of this is (Robbins 1983):

$$\mathring{m}^i_{x,4} = \hat{m}^i_{x,4} + S^i_{x,4}(\bar{m}^i_{x,4} - \hat{m}^i_{x,4})$$

Again, it's empirical because $m_{x,4}$ it is replaced by $\bar{m}_{x,4}$, the weighted average of the small areas. The factor $S^i_{x,4}$ ("shrinkage") is the ratio between the expected sample variance of the estimator in the smaller area and the unconditional variance of the estimator, which ends up being:

$$S^i_{x,4} = \frac{V_m(m^i_{x,4})}{V_m(m^i_{x,4}) + \frac{m_{x,4}}{N^i_{x,4}}}$$

From other point of view, the ratio between the variance of the smaller area with respect to the sum of the total variance (of the smaller and larger area), in tune with a classical variance analysis (ANOVA). Following this reasoning, in a context of extreme homogeneity, a very small minor area could be characterized based on the estimation of the largest area ($S^i_{x,4} \cong 1$). Moreover, areas of high population weight will take values close to those observed ($S^i_{x,4} \cong 0$). In the middle of these extremes, the linearly function combines the estimation of the largest area with respect to that of the smaller area, with the weighting between both being such that the jurisdictions with small population (the extreme case would be one that does not present events in any age category), will be closer to the greater area than those with higher population.

Longford (1999) extended this idea to vectors of random variables from smaller areas, with

the possibility of taking advantage of the correlation between them. In our case, if the mortality rate of the age group between x and x + 4 of area i is greater than that of area j, that the correlation is high would imply that in a contiguous age the same happens with more chance. The calculations followed in this paper based on this idea, follow the methodological development of Assuncao [(2005): 543-544) which estimates the parameters by the moment method for the case of fertility rates in Brazil.

References

Assunção, Renato M., Carl P. Schmertmann, Joseph E. Potter, and Suzana M. Cavenaghi. 2005. "Empirical Bayes Estimation of Demographic Schedules for Small Areas." *Demography* 42 (3). Springer: 537–58. http://www.jstor.org/stable/4147361.

Efron, Bradley, and Carl Morris. 1972. "Empirical Bayes on Vector Observations: An Extension of Stein's Method." *Biometrika* 59 (2). [Oxford University Press, Biometrika Trust]: 335–47. http://www.jstor.org/stable/2334578.

Freire, Queiroz, F. H. M. d. A. 2015. "Mortality Estimates and Construction of Life Tables for Small Areas in Brazil, 2010."

Grushka, Carlos O. 2014. "Casi Un Siglo Y Medio de Mortalidad En La Argentina..." *Revista Latinoamericana de Población* 8 (15, Julio/Diciembre): 93–118. http://revistarelap.org/ojs/index.php/relap/article/view/14/13.

Lima, Everton Emanuel Campos de, and Bernardo Lanza Queiroz. 2014. "Evolution of the Deaths Registry System in Brazil: Associations with Changes in the Mortality Profile, Under-Registration of Death Counts, and Ill-Defined Causes of Death." *Cadernos de Saúde Pública* 30 (8): 1721–30. https://doi.org/https://dx.doi.org/10.1590/0102-311X00131113.

Lima, Everton Emanuel Campos de, Bernardo Lanza Queiroz, and Diana Oya Sawyer. 2014. "Método de Estimação de Grau de Cobertura Em Pequenas áreas: Uma Aplicação Nas Microrregiões Mineiras." *Cadernos Saúde Coletiva* 22 (4): 409–18. https://doi.org/https://dx.doi.org/10.1590/1414-462X201400040015.

Longford, N. T. 1999. "Multivariate Shrinkage Estimation of Small Area Means and Proportions." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162 (2): 227–45. https://doi.org/10.1111/1467-985X.00132.

Marshall, Roger J. 1991. "Mapping Disease and Mortality Rates Using Empirical Bayes Estimators." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 40 (2). [Wiley, Royal Statistical Society]: 283–94. http://www.jstor.org/stable/2347593.

Müller, M. M., M. S.; Accinelli. 1980. "Un Hecho Inquietante: La Evolución Reciente de La Mortalidad En La Argentina." *Cuaderno Del CENEP* 17. http://201.231.155.7/wwwisis/bv/cuadernos%20cenep/CUAD%2017.pdf.

Queiroz, B. L., Flávio Henrique Miranda de Araujo Freire, Marcos Roberto Gonzaga, and Everton Emanuel Campos de Lima. 2017. "Completeness of Death-Count Coverage and Adult

Mortality (45q15) for Brazilian States from 1980 to 2010." *Revista Brasileira de Epidemiologia* 20 (Suppl. 1): 21–33. https://doi.org/https://dx.doi.org/10.1590/1980-54972017000500033.

Recchini de Lattes, Z. L., and Alfredo E. Lattes. 1975. *La Población de Argentina*. Book. C.I.C.R.E.D. Series. Buenos Aires: Talleres Gráficos Zlotopioro, INDEC.

Robbins, Herbert. 1983. "Some Thoughts on Empirical Bayes Estimation." *The Annals of Statistics* 11 (3). Institute of Mathematical Statistics: 713–23. http://www.jstor.org/stable/2240635.

Schmertmann, Carl P., and Marcos R. Gonzaga. 2018. "Bayesian Estimation of Age-Specific Mortality and Life Expectancy for Small Areas with Defective Vital Records." *Demography* 55 (4): 1363–88. https://doi.org/10.1007/s13524-018-0695-2.

Setel, Philip W., Sarah B. Macfarlane, Simon Szreter, Lene Mikkelsen, Prabhat Jha, Susan Stout, and Carla AbouZahr. 2007. "A Scandal of Invisibility: Making Everyone Count by Counting Everyone." *The Lancet* 370 (9598): 1569–77. https://doi.org/10.1016/S0140-6736(07)61307-5.