Extended Abstract Submitted to the Annual Meeting of the Population Association of America - 2013

Adult Mortality Estimates for Small Areas in Brazil, 1980-2010

Everton Emanuel Campos de Lima
Department of Demography – Universidade Federal de Minas Gerais
(everton@cedeplar.ufmg.br)

Bernardo Lanza Queiroz
Department of Demography – Universidade Federal de Minas Gerais
(lanza@cedeplar.ufmg.br)
Introduction

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).

In this context, the 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. This paper aims to study the evolution of adult mortality in small areas in Brazil from 1980 to 2010. We use 45q15 as a
summary measure of adult mortality across small areas in Brazil. In order to produce the estimates, we propose a methodological approach that combines the death distribution methods (DDM) to indirect standardization in order to produce more reliable estimates of adult mortality for small areas. We focus on adult mortality because there are much more studies on infant and child mortality for sub-national population using indirect demographic techniques (Souza, Hill and Dal Poz, 2010; Castro and Simões, 2009). The studies on child and infant mortality are also showing a convergence of levels across regions in Brazil, thus the main changes in life expectancy in recent years should be explained by changes and variation in adult mortality levels. In addition to that, the combination of reliable estimates of child (5q0) and adult mortality (45q15) can be used to estimate complete life-tables for small areas in Brazil with more precision using the method proposed by Wilmoth, et al (2012).

A preliminary study on estimating adult mortality in Brazil was presented by Lima et.al (2012). In that study, the authors focused on the state of Minas Gerais (about 10% of the Brazilian population), produced estimates for both sexes combined, and in only one point in time. In this paper, we extend the estimates for the whole country (137 small areas), estimating mortality for males and females separately and covering the period from 1980 to 2010. The paper will also provide spatiotemporal estimates useful for future research, especially to analyze trends in adult mortality in Brazil and investigate how changes in socioeconomic conditions are related to changes in adult mortality.

The problem of small areas mortality rates

The availability of information on population, vital events and other such as health and geo-referenced data, along with statistical and computational advances, allows researchers to study and investigate demographic conditions in small areas. Such analyzes are important because they allow researchers to investigate environmental and behavioral aspects of disease, access to health care, and better understand the socioeconomic determinants of mortality and morbidity in these areas (Ferguson et al., 2004). However, such studies often suffer with problems related to instability in the estimates of rates in least-populated areas (Assunção et al., 2005; Pollard, 1970). 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 estimation, 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).

In the case of mortality data and estimates, Lima and Queiroz (2011) showed how the degree of coverage of deaths in the micro health regions of the state of Minas Gerais is characterized by a strong spatial pattern between 1980 and 2007, with high levels of death counts under-registration in areas Northern and Northeastern of the state. The same spatial pattern is observed in the regions of Brazil between 1980 and 2006 (Lima and Queiroz, 2011). Cavalini and Pocé de Leon (2007) produced estimates of death counts under-coverage for the country, using empirical bayes, at similar levels of others using traditional demographic methods (Agostinho and Queiroz, 2008). Despite the concise results, the authors emphasize the fragility of the estimates of mortality under-registration, given the limitations that traditional methods of demographic of mortality estimate for sub-national population groups (Hill et al., 2009; Timaeus, 2001). Thus, the search for new means of evaluates mortality data and produce better estimates of mortality schedules emerges as a major issue in Brazilian demography.

**Data and methods**

We make use of the mortality database available at the Brazilian Ministry of Health Database - Datasus (França, et.al, 2008; Ministério da Saúde, 2012). The data are collected by age, sex and causes of death at the municipality level. Population by age and sex comes from national household census conducted by the National Statistics Office in 1980, 1991, 2000 and 2010. We aggregated municipalities by comparable small areas, using the National Statistics Office (IBGE) definition of comparable mesoregions. These regions are constructed utilizing regional and socioeconomic similarities. The regions only
serve a statistical purpose; therefore, they do not represent a political or administrative entity. The main advantage of working with these comparable areas is that they have not changed their boundaries over the period of analysis. Thus, we are able to follow and study 137 small areas between 1980 and 2010.

**Mortality Coverage Evaluation**

To evaluate the coverage of reported deaths we use traditional demographic methods, called Death Distribution Methods – DDM henceforth (Hill, You and Choi, 2009; Dorrington, 2012a; 2012b). The DDM are commonly used to estimate adult mortality in a non-stable population and analyze mortality data quality in intercensal periods (Timeaus, 1991; Hill et al, 2005; Hill, You and Choi, 2009). They make several strong assumptions: 1) that the population is closed to migration; 2) that the completeness of recording of deaths and population are constant by age; and 3) that ages of the living and the dead are reported without error. There is a large body of literature on the methods, for reasons of space they will not be discussed in detail here.

The assumption that the population is closed to migration is important to Brazilian regions, since the country is marked by significant migration flows between its regions (Barbieri et al, 2010). The DDM method uses information on deaths and growth rates accumulated above a series of ages x. If there is some age x above which net migration is negligible, the performance of the methods above that age will be unaffected (Hill, You and Choi, 2009; Murray, et.al, 2009; Dorrington, 2012). We use the age range 30+ to 65+ as suggested elsewhere, (Hill et al, 2009; Queiroz, 2011) to avoid possible problems regarding migration and to overcome limitations as a result of old age reporting errors (referring to age declaration). The adjustment factor used is the average of the three methods (GGB, SEG and SEG-adj) and using the age range 30+ to 65+.

**Combining indirect standardization with death distribution methods**

The estimates for adult mortality in Brazil followed the methodology applied by Lima et al (2012) for Minas Gerais, Brazil, to test the viability of the method. The methodology consists in a combination of indirect standardization with death distribution methods (DDM). In order to evaluate the mortality data and to estimate adult deaths rates in small areas, the authors first applied an indirect standardization as smoothing method to get more stable mortality age schedules for small areas in Minas Gerais, Brazil, between 1991 and 2000. Afterwards, they applied the DDM to correct the mortality levels of these small areas and subsequently estimate adult mortality.
The indirect standardization is a useful technique since it allows not only to compare rate levels between different populations, but also to estimate vital rates schedules in population which does not have reliable data information (Preston et al., 2001). The method takes a function (or a set of age-specific rates) from a population that, a-priori, is considered similar to the study population (Lima et al., 2012). In that study, the function is originally taken from the bigger region from where the small belongs and it is assumed that the pattern of mortality for both populations, small and big, are very similar during the period of analyses. The principal is similar to the Bayesian idea of minimize rates fluctuation errors by using the vital record information of the neighborhood (Marshall, 1991). The indirect standardization is as follows:

\[ n \bar{\sigma}_{x_i}^t = nN_{x_i}^t \times n \bar{m}_{x_i}^t \]  

Where:

- \( n \bar{\sigma}_{x_i}^t \): Number of expected deaths in the small area \( i \) in time \( t \) between the ages \( x \) e \( x+n \);
- \( nN_{x_i}^t \): Population in small area \( i \) in time \( t \) between ages \( x \) e \( x+n \);
- \( n \bar{m}_{x_j}^t \): Age-specific mortality rates between ages \( x \) e \( x+n \) in big area \( j \) in time \( t \).

Thus, using the age-specific mortality rates of bigger areas, they estimate the expected number of deaths for the smaller areas between the years 1991 and 2000. In the second step, they analyze the quality of mortality data for small areas (comparing the new observed death counts to the death counts expected by demographic changes) of death records and estimated measures of adult mortality for the same period (Lima et al. 2012).

Figure 1 shows the results from Lima et al (2012) for Minas Gerais. The application of the proposed methodology involves an evaluation of mortality schedules of small areas in the given period. The figure shows the age-specific mortality rates with and without standardization. In the most regions the observed rates present a very unstable mortality age structure. These variations are even more pronounced when compared with the age-specific rates estimated by indirect standardization. It should be noted, however, that the levels of both schedules are relatively similar, i.e. the proposal method is correcting the age-specific mortality fluctuations, but it does not to impose a very different levels of mortality.
Figure 1: Age-specific mortality rates, estimated directly and through indirect standardization. Small regions of Health of Minas Gerais, 1991-2000

Source: Lima et al. (2012).
Estimating adult mortality in small areas

Table 1 presents estimates of adult mortality, 45q15 (probability of dying between ages 15 and 59), for each region in Minas Gerais. The results are also extracted from Lima et al. (2012).

We present the results on the basis of four (4) different estimates: directly - without correction of the data; applying the methods of distribution of deaths; via indirect standardization - uncorrected data; and applying the methods of distribution of deaths after standardization of death rates (combination of methods).

The results show that different methods give different estimates of adult mortality. However, the differences are on average very small, but they are larger for the least developed areas of the state where worse data quality is more likely to occur. The Northeast region of Minas Gerais is also the area that presents the greatest limitations of data for the larger area (used as the standard) indicating a worse quality of vital information for the small which it belongs. In the localities where they find more stable mortality rates by age for both small and larger areas, and where data quality in larger areas are of good quality, the estimates are very similar.

We estimate adult mortality, 45q15, in 15% for the state when using direct estimation, without correction. The second set of estimates, after adjusting the method for under-registration of death counts, indicates adult mortality of 19%. The adjustment factor used is the average of the three methods (GGB, SEG and SEG-adj) and using the age range 30+ to 65+. Estimates based on two-stage method (indirect standardization + DDM) produce estimates of adult mortality a bit higher (45q15 of 20%) than when it is applied only DDM methods, but with a much more stable age-profile than when we estimate mortality rates directly from the data.

Some specific cases deserve further attention. In particular, when there is no correction applied, the small area of Mantena, Águas Formosas, Pedra Paraiso and Itaobim present a very low adult mortality levels from 1991 to 2000. Since those regions are located in the poorest areas of Minas Gerais, we expected higher levels of adult mortality. A visual inspection of the age profile of mortality indicates that the data is very limited and rates are very unstable. These elements can help explain the obtained mortality estimates. After applying the different correction methods, we obtain a significant increase in levels of adult mortality, which normally could be expected in these locations.

The difficulty is, then, to define the correct level of mortality in regions where the data received major adjustments. On the one side, we have locations where estimates of
the different methods are very close, which may be a good indication of the quality of the estimates and/or good quality data. On the other side, there are regions with a very high variation of estimated 45q15. For example, the case of Pedra Azul drives our attention. In this location the 45q15 observed is 0.14. The estimate cumulative probability using the two-stage method is 0.22. Thus, additional analyzes are important to get the best estimate of mortality in these localities. For example, the analysis of the relationship between infant mortality and adult mortality, which according Wilmoth et al. (2012), seems to be quite stable over time and across countries, may be an interesting possibility as a complementary study.

Table 1: Estimates of Adult Mortality. Small regions of Minas Gerais, 1991-2000

<table>
<thead>
<tr>
<th>Small area</th>
<th>Big area where the small location belongs</th>
<th>45q15</th>
<th>No mortality correction</th>
<th>DDM</th>
<th>No mortality correction</th>
<th>DDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Três Corações</td>
<td>South (Alfenas/ Pouso Alegre/ Poços De Caldas/ Passos/ Varginha)</td>
<td>0.18</td>
<td>0.19</td>
<td>0.17</td>
<td>0.17</td>
<td>0.18</td>
</tr>
<tr>
<td>Guanhães</td>
<td>Center (Belo Horizonte)</td>
<td>0.15</td>
<td>0.19</td>
<td>0.19</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>M Novas/ Turmalina/Capelinha</td>
<td>Jequitinhonha (Diamantina)</td>
<td>0.19</td>
<td>0.21</td>
<td>0.20</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>Bom Despacho</td>
<td>West (Divinópolis)</td>
<td>0.18</td>
<td>0.18</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Itaúna</td>
<td>West (Divinópolis)</td>
<td>0.19</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Mantena</td>
<td>East (Governador Valadares/ Ipatinga)</td>
<td>0.12</td>
<td>0.15</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Resplendor</td>
<td>East (Governador Valadares/ Ipatinga)</td>
<td>0.16</td>
<td>0.18</td>
<td>0.17</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Além Paraíba</td>
<td>Southeast (Juiz De Fora)</td>
<td>0.14</td>
<td>0.13</td>
<td>0.18</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>Santos Dumont</td>
<td>Southeast (Juiz De Fora)</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td>São João</td>
<td>Southeast (Juiz De Fora)</td>
<td>0.17</td>
<td>0.18</td>
<td>0.19</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Nepomuceno/Bicas</td>
<td>Region Norte De Minas (Montes Claros)</td>
<td>0.18</td>
<td>0.20</td>
<td>0.18</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Coração de Jesus</td>
<td>Region Norte De Minas (Montes Claros)</td>
<td>0.16</td>
<td>0.21</td>
<td>0.18</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Francisco Sá</td>
<td>Region Norte De Minas (Montes Claros)</td>
<td>0.16</td>
<td>0.21</td>
<td>0.18</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Águas Formosas</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.10</td>
<td>0.15</td>
<td>0.18</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Araçuí</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.17</td>
<td>0.25</td>
<td>0.17</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>Itaobim</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.11</td>
<td>0.19</td>
<td>0.17</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Nanuque</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.16</td>
<td>0.25</td>
<td>0.17</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Padre Paraíso</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.10</td>
<td>0.20</td>
<td>0.18</td>
<td>0.19</td>
<td>0.19</td>
</tr>
<tr>
<td>Pedra Azul</td>
<td>Northeast (Teófilo Ottoni)</td>
<td>0.14</td>
<td>0.17</td>
<td>0.18</td>
<td>0.22</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Mean                     | 0.15  | 0.19                    | 0.18| 0.20                    | 0.20|

No mortality correction¹: the estimates are based on age-specific mortality rates without correction.
No mortality correction²: the estimates are based on age-specific mortality rates indirect standardized according to the mortality schedules of big areas.
DDM¹: Death Distribution Method.
DDM²: Death Distribution Method estimate after an indirect standardized.
Expected results: analyzing data quality and estimating adult mortality for small areas in Brazil.

Spatiotemporal evaluation of mortality coverage

We first show the spatiotemporal evolution of the completeness of death counts from 1980-1991 to 2000-2010 for both sexes separately. These estimates are based on DDM, without the indirect standardization. This will be the next step of this paper.

The completeness of death counts considerably increased for both sexes. From average 80% in 1980-1991 to average 95% in 2000-2010, but a large regional variation still persists. In the intercensal period of 1980-1991 coverage ranged from 6% to 100%, for both sexes, and only the mesoregions in the South and Southeast presented reasonable data quality. In the intercensal period of 1991-2000 data quality improved to other parts of the country, except in many parts of the Northeast and North.

In the period 2000-2010, the degree of coverage varies from 17% to 32% in the worst areas of the Northeast and North to 100% in several mesoregions of South and Southeast. The inverse of estimated coefficient can also be used to adjust the number of deaths by age and provide better estimates of mortality and life expectancy.

The estimated coefficients indicates that a large number of regions in Brazil does not collect complete information on the number of deaths that occurred in the region, this leads to overestimation of life expectancy and might mislead public health policies. It is also important to notice that the levels of completeness of death counts are generally higher among males than females.
Figure 2: Completeness of death counts coverage, Death Distribution Methods, males and females, 1980-2010

Estimation of adult mortality in small areas

We present some results of mortality schedules and cumulative deaths probability of two selected mesoregions, namely the mesoregion of Sudoeste Amazonense and Norte do Amapá. These regions are located in the North Amazon area where the completeness of deaths counts is one of the worse of the country.

Figure 3 and Figure 4 shows the results for males and females, respectively, from the intercensal period 1980-1991 to 2000-2010. The figures show that age-specific mortality rates have a very erratic pattern making it impossible to produce adequate estimates of adult mortality. For some ages, in particular areas, the vital records indicated that not a single death occurred at that age. This could happen for many reasons. For instance, it can be a problem of death count under-registration, problem of small numbers or a real feature of the data. In any case, it is important to find ways to produce more reliable estimates of adult mortality for small areas in Brazil.

Table 2 presents preliminary estimates of adult mortality, 45q15, for males and females in the same regions presented in the Figures 3 and 4. We only show results based on direct estimation and after adjusting the data for incompleteness of death counts coverage, using DDM methods. The adjustment factor used is the average of the three methods (GGB, SEG and SEG-adj) and using the age range 30+ to 65+. The first important results is a clear problem of incompleteness of death counts coverage (as we saw in Figure 2), since all estimates of 45q15 obtained after correcting the data are much greater than the direct ones. However, those estimates should be considered with caution, since it is clear that the age profile of mortality has severe problems that might affect the cumulative probabilities estimates. For example, some estimates are presenting a variation in 45q15 of more than 70% between methods. Thus, there is a clear problem with the mortality schedules of those areas which needs more investigation.

The next step of the paper is to produce estimates of adult mortality using the two-stage method (indirect standardization + DDM) presented before. This time, we will use states mortality schedules as the standard to adjust age-specific mortality rates in the mesos-regions.
Figure 3: Age-specific mortality rates estimated directly, males, Brazil, mesoregions, 1980-2010.

Figure 4: Age-specific mortality rates estimated directly, females, Brazil, mesoregions, 1980-2010.

## Table 2: Estimates of Adult Mortality, mesoregions of Brazil, 1980-2010

(Female and Male Separately)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45q15 of Female</td>
<td>45q15 of Female</td>
<td>45q15 of Female</td>
</tr>
<tr>
<td></td>
<td>No mortality</td>
<td>No mortality</td>
<td>No mortality</td>
</tr>
<tr>
<td></td>
<td>correction* DDM†</td>
<td>correction* DDM†</td>
<td>correction* DDM†</td>
</tr>
<tr>
<td>Sudoeste Amazonense</td>
<td>0.01</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td>Norte do Amapá</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>0.12</td>
<td>0.61</td>
<td>0.22</td>
</tr>
</tbody>
</table>

*No mortality correction: the estimates are based on age-specific mortality rates without correction.
†DDM: Death Distribution Method. We used an average of the 3 methods (GGB, SEG and SEG-adj), estimating the completeness of death counts using the age range 30+ to 65+.

### References


