!
!

National and subnational experience with estimating the extent and trend in completeness of registration of deaths in Brazil[1]

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

Everton E. Campos de Lima
NEPO - Unicamp
everton@nepo.unicamp.br

Abstract

In this paper, we analyze the evolution of the completeness of death counts coverage in Brazil and its regions since 1980. We review a series of studies on the quality of mortality registration for the country, states and small areas, compare and contrast different approaches and results. We also investigate the quality of the 2010 Census

P a a
C a

[1] Research presented in this paper builds on collaborative work with Everton Lima (Unicamp), Flávio

!

!
!
file is transferred to the national office every 3 months. Population by age and sex, at the local level, comes from the Brazilian Censuses (1980, 1991, 2000 and 2010).

We use the information available in the Brazilian Census 2010 on the household deaths that occurred in the last 12 months. The respondent also informs sex and age of the person who died during the reference period. The inclusion of such information information is part of the suggestions made by the United Nations Principles and Recommendations for Population and Housing Censuses (1997) for the 2000 and 2010 census round. Household deaths should be evaluated very carefully because it is subject to four main problems: a) sub-enumeration of deaths due to breakage of households following the death of a resident or due to high concentration of residents in single person households; b) young people who migrate from their original homes can be registered as belonging to more than one domicile; c) errors in the reference period for that the definition of the period must be as clear as possible and d) possible random and significant fluctuations in mortality affecting the estimate of interest.

2.3. Death Distribution Methods

The demographic literature has developed several methods based upon equations of population dynamics to evaluate the coverage of reported deaths relative to populations. The death distribution methods (DDM) are commonly used to estimate adult mortality in a non-stable population. A stable population is one in which the birth and death rates are unchanging over a long period of time. The DDM methods compare the distribution of deaths by age with the age distribution of the living and provide age

!
!

Where N'(x) is the number of persons who reach the exact age x in the period, N(x+) is the number of persons at exact age x and over, r(x+) is the population growth rate, k1 and k2 are the relative coverage of the enumerated population in two censuses, C is the degree of completeness of death records over the period, D'(x+) is the observed number of deaths of people with x or more years of age and t is the interval corresponding to the intercensal period

This identity holds for open-ended age segments x+, and in a closed population the only entries are through birthdays at age x. The "birth" rate x+ minus the growth rate x+ thus provides a residual estimate of the death rate x+. If the residual estimate can be calculated from population data from two population censuses and compared to a direct estimate using the recorded deaths, the completeness of death recording relative to population recording can be estimated. The method has a few strong and key assumptions: the population is closed to migration; the completeness of recording of deaths is constant by age; the completeness of recording of the population is constant by age; and ages of the living and dead are reported without error.

The death distribution method has also an important limitation. Since they compare a distribution of deaths to an intercensal population, they estimate intercensal completeness of mortality estimates, and not the completeness at the beginning or end of the intercensal period. This matter is of particular importance when a distribution of deaths comes from data from the latest census or when vital registration is available for recent years,. The assumption of the GGB method that the population is closed to migration is also of importance to Brazil and its regions. The GGB method uses information on deaths and growth rates cumulated 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. The intercept and slope of the GGB method were obtained by orthogonal regression to points for the ages 35+ to 75+.

The Bennett and Horiuchi method, SEG, uses specific growth rates by age for converting an age distribution of deaths into an age distribution of a population. Once the observed deaths from a given age x in a population is equal to the population of age x, adjusted by the rate of population growth by age range, we have the deaths of a population of age x+ that provide an estimate of the population on that age x. The extent of death registration coverage is given by the ratio of deaths estimated by the population

!

!

!
!

measures of interest. This instability is even worse when sub-national groups are disaggregated by age and sex (Assunção et al. 2005). 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 show 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) showed how Empirical Bayes was effective in the case of Brazilian municipalities. 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).

3. Results

Table 1 shows the trends in completeness of death counts coverage in Brazil, for males and females, using different methods and age ranges to estimate completeness and levels of adult mortality. Completeness improved systematically for the country since 1980, but with important differences between males and females. In general, completeness of death counts coverage is better for males than for females. The results indicate that coverage in Brazil increase, for males, from 84% in 1970-1980 to 95% in 1991-2000, we will discuss 2000-2010 in detail later. For females, we find that completeness improved from 79% to 90%. One important message from Table 1 is that different methods and different age ranges used to produce estimates can produce different results. Since the literature does not specify gold-standard method, it is important to present the results using different approaches.

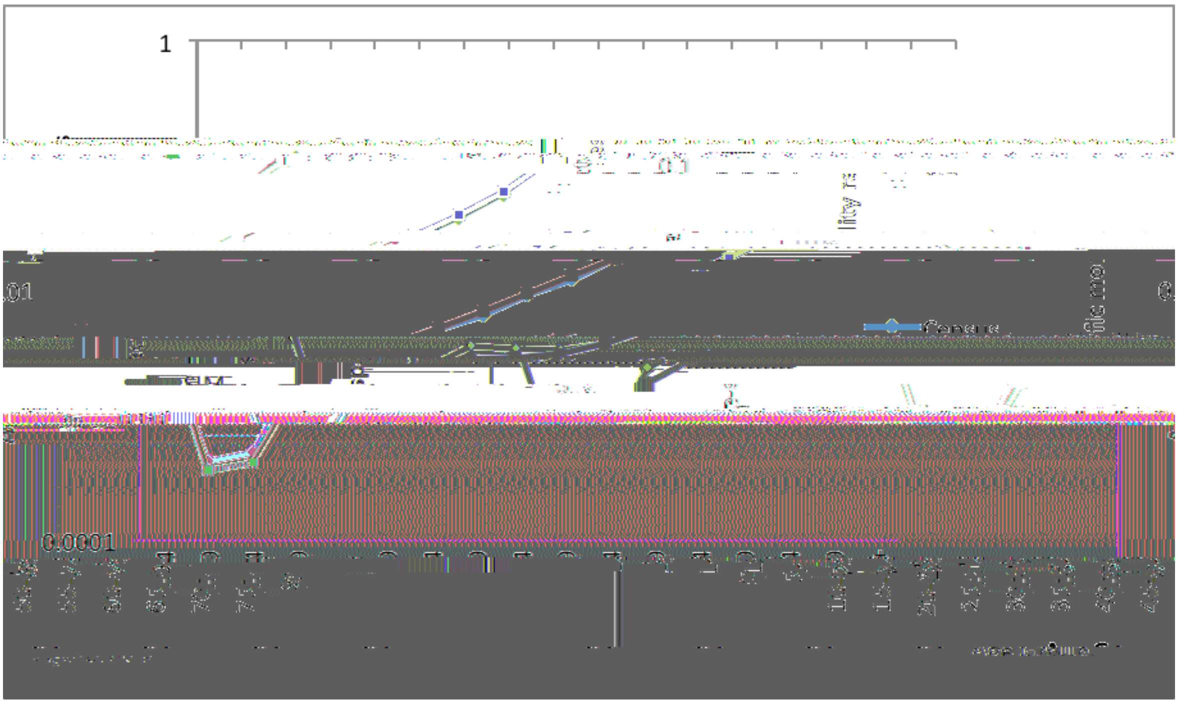
!
!
the degree of coverage of mortality data. Queiroz and Swayer estimated completeness age intervals from 5+ to 65+ for the GGB method and 15+ to 65+ for the SEG method. The adult mortality estimates use completeness from the GGB method, but they poin that results are quite similar.

Tabela 2 – Completeness of Death Counts Coverage (SIM-Datasus) and Household Deaths (Census), 2010

	Males Census	Males SIM - Datasus	Females Census	Females SIM -Datasus
GGB				
k1/k2	0,9926	0,9919	0,9864	0,9890

!

!



Source: Brazilian Population (2010) e SIM-Datasus (2010)

We perform a simple exercise to compare estimates age specific probabilities of death produced by IBGE (official estimates) to an application of Wilmoth et.al (2012) model life-table. The method does not produce an estimate of completeness of death counts coverage, but allows one to estimate full age profiles of mortality with limited or defective data.

Equation 1 formalizes the model:

(Equation 1) , where

$\log(m_x)$ is the age specific mortality rate,

$h = \log({}_5q_0)$ reflects the level of child mortality,

!
!

The main disadvantage is that the constants in the model were constructed using countries in the Human Mortality Database and the experience of those countries might not reflect the experience of less developed countries. However, the method is very flexible and one could re-estimate the constants by adding other countries to the database. The main issue is that those additional countries might have data limitation and mortality age-profiles might have been estimated using indirect methods or model life-tables. Figure 2 compares the official estimates for Brazil in 2000 (IBGE) to the application of the log-quad model using one and two parameters. We use 5q0 and 45q15 from our estimates presented in the previous section. The main issue is the difference between the observed level in adult mortality by IBGE and using the log-quad methods. IBGE estimates 45q15 in 0.2584 whereas one-parameter estimates is 0.2052 and two-parameter estimates is 0.2460.

Figura 2 – Age specific probabilities of death between ages 15 and 60, 45q15, Males, Brazil, 2000.

!

!

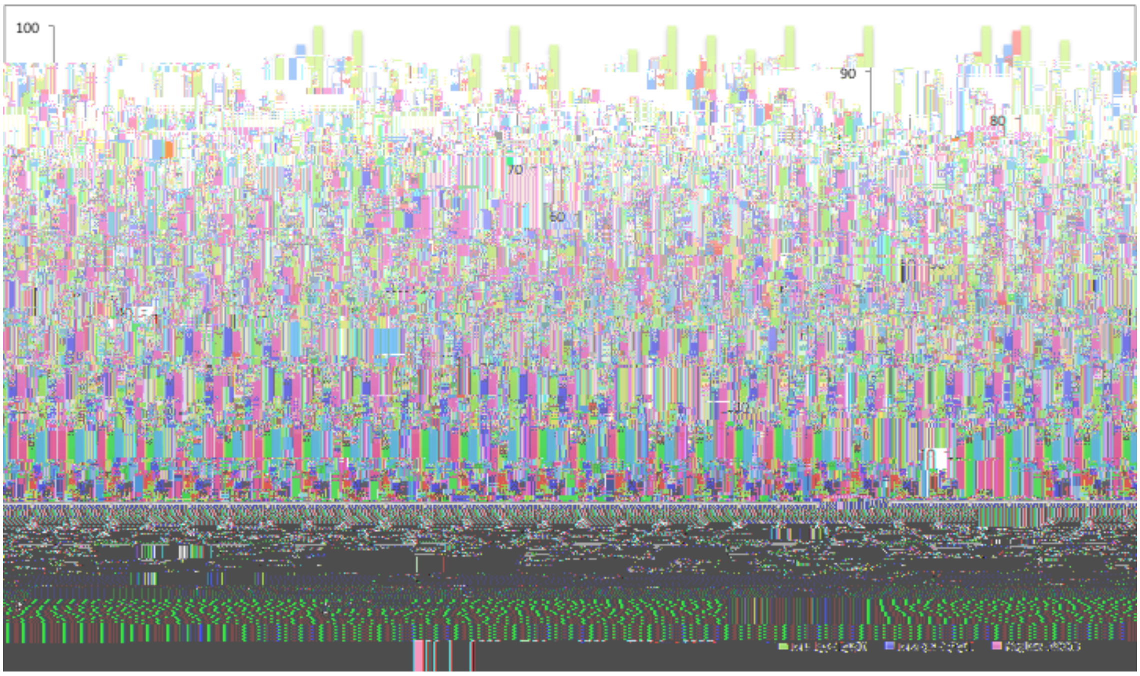
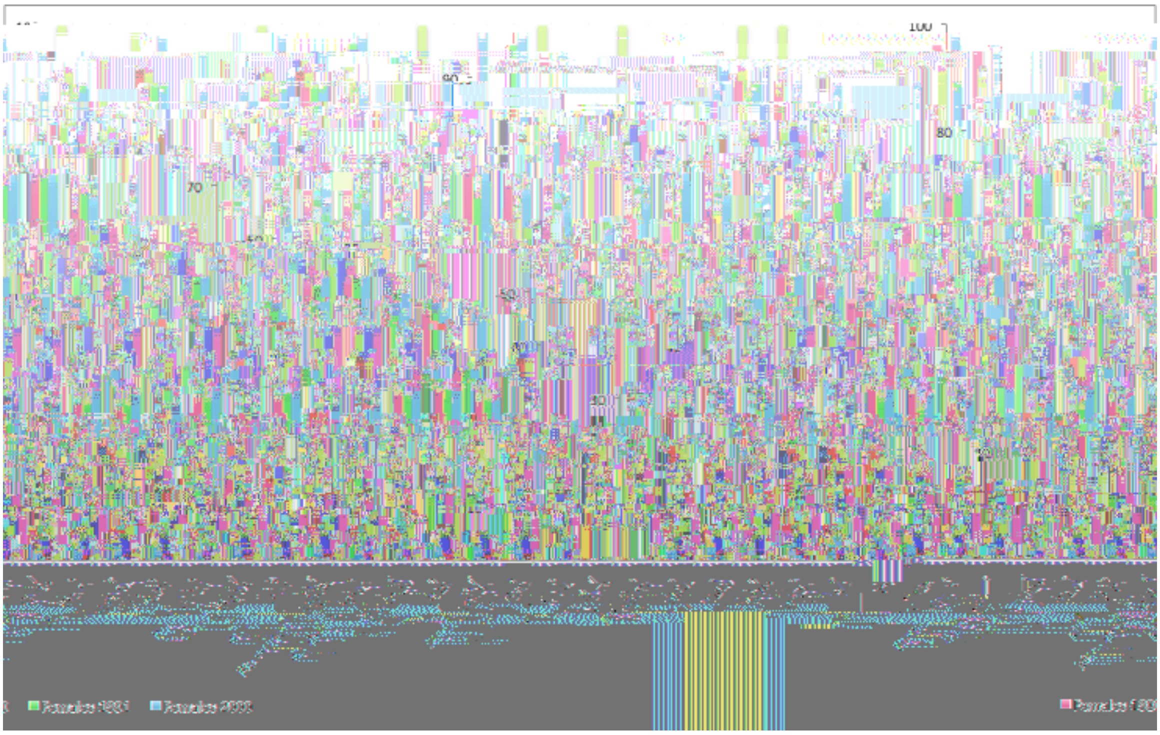


Figure 4 – Completeness of Death Counts Coverage, average of Death Distribution Methods, Brazilian States, Females, 1980 to 2010.



!

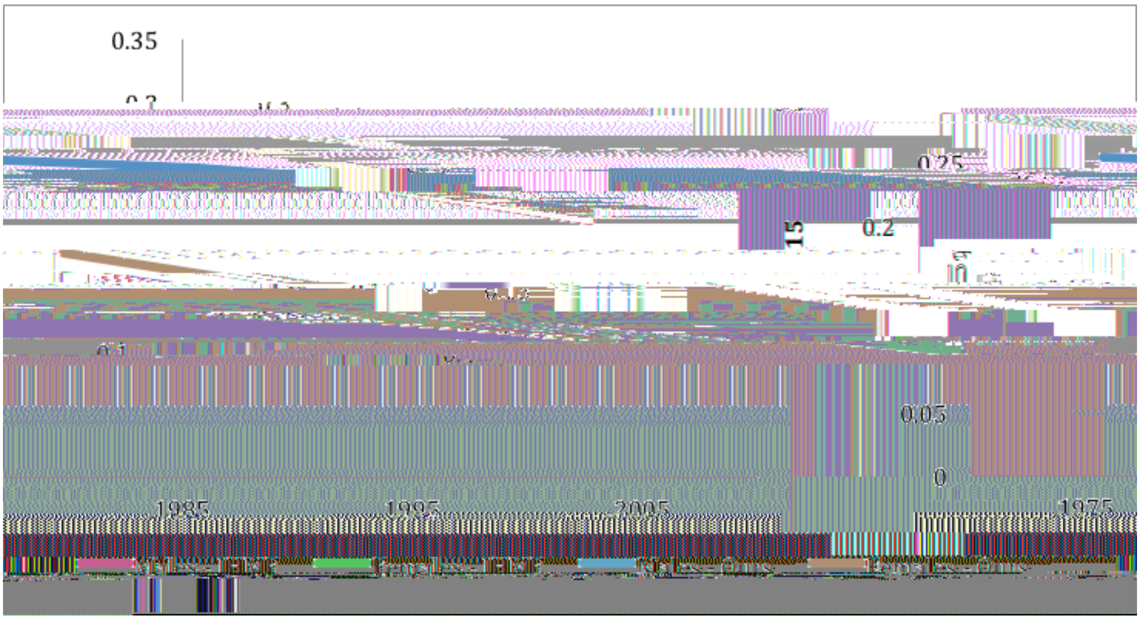
!

!

!

Burden of Diseases project. Levels of mortality estimated by IHME are always lower than obtain by our researcher group, for both males and females. More important, it seems that coverage level estimated by IHME differs in substantive magnitude from our estimates for the earlier period of analysis, for example, their estimates in the 1980's are higher than obtain using observed data from Datasus. More important, we see that estimates show a very different trend in the levels of adult mortality over the last three decades.

Figure 5 – Comparison between adult mortality estimates from IHME and our estimates using DDM, Brazil 1970-2010 (mid-census years)

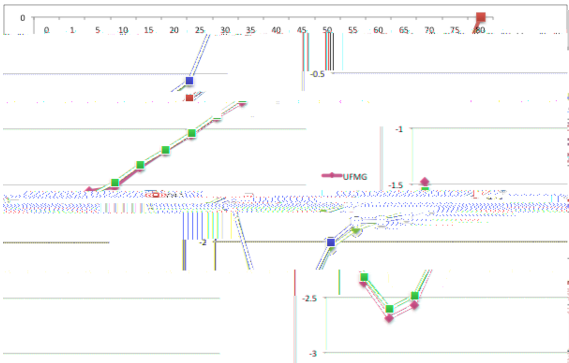


We also find signiticative differences at the state level. Our estimates indicate a

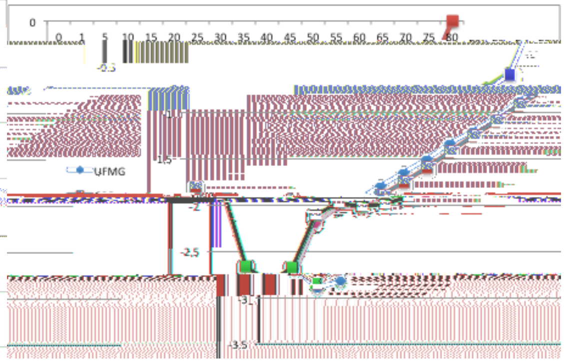
!

!

Maranhão



São Paulo



3.3. Evolution of Death Counts Coverage by Small Areas

In a recent paper, Queiroz, et.al (2016)[2] estimated mortality levels for mesoregions in Brazil from 1980 to 2010. The paper uses three-step procedure to estimate mortality in small areas when data is defective. First, they apply a standardization technique to smooth rates in small areas. They standardize mortality rates from mesoregions (smaller areas), by sex, using state level data as standards Second, they obtained measures of completeness of death counts coverage through the Death Distribution Methods. Finally, we perform a bottom-up adjustment to make sure that adding up death counts at the local level we obtain the total number of deaths in each state and the country.

The original data is available at the municipality level. The main limitation in using city level data in Brazil is that the number and composition of cities change over time. In 1980, there were 3974 municipalities and in 2010 there were 5565. To avoid problems using this information, Queiroz and colleagues (2016) aggregated municipalities by comparable small areas, using the IBGE definition of comparable mesoregions. Mesoregions are stable and comparable over the period of analysis and it is possible to follow the same 137 regions from 1980 to 2010 (and more recent years).

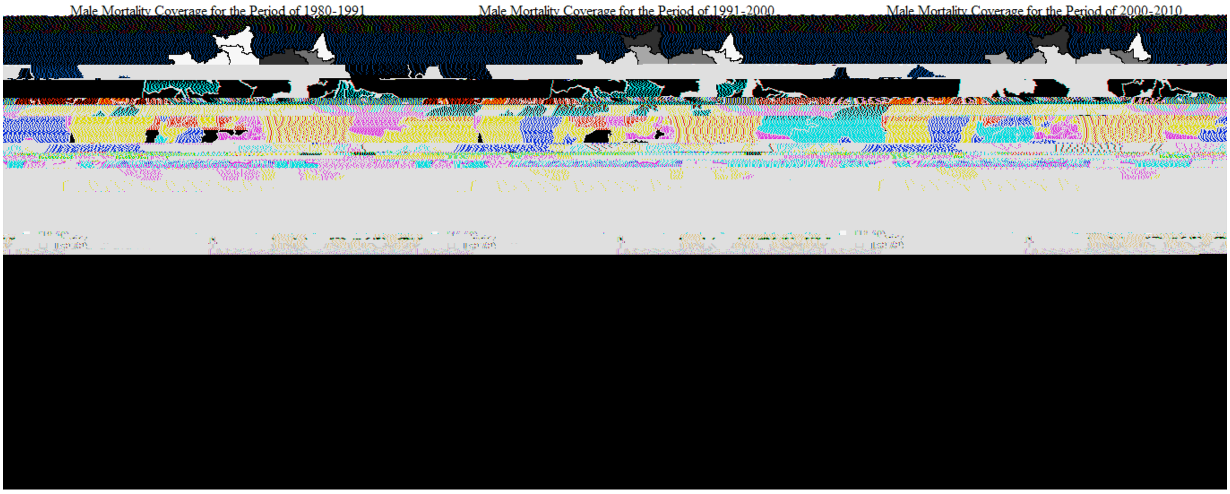
Map 1 and Map 2, obtain from Queiroz and colleagues (2016), depicts the evolution of completeness of death counts in Brazil by sex and mesoregions from 1980 to 2010. We show before that completeness of death counts coverage is improving overtime in

[#]! , -./!01!01!2.-3.455!67!894!.454:.;9!3.-<2!89:8!=>?!:1@!AAB>!:.4!2:.8C!D94!2:24.!05!01!E01:F!2.42:.:80-1! 8-!64!5<6G0884@!8-!:!5;04180E0;!H-<.1:FC!I8!05!2:.8!-E!:!F:.34.!.454:.;9!2.-H4;8!-1!4580G:8013!G-.8:F087!01! 5G:FFJ:.4:5!01!=.:K0F!E<1@4@!67!BLMNC!O45<F85P!@:8:!:1@!;-@45!:.4!:Q:0F:6F4!<2-1!.4N<458C!R!S46J5084!S0FF! 64!.4F4:54@!S9414Q4.!894!2:24.!05!:;;4284@!8-!2<6F0;:80-1C!!

!

!
!
the country, but there are still significant regional variation. In the South and Southeast regions have better coverage of deaths counts than others regions in 1980 was over 90% reaching almost 100% in the more recent years. In the North and Northern parts of the country, completeness, is improving rapidly over the last decades (from 60% to 90% in the period).

Map 1 – Completeness of Death Counts Coverage, Males, Brazil, 1980-2010



Source: Queiroz, et.al, (2016)

!

!
!

able to produce complete and smooth curves, however, a question that remains is which modelgives reliable estimates of mortalityin those small areas, and if those estimates are reasonable given the knowledge about the evolution of mortality in the country.

The curves estimated by the models do not vary in terms of structure, but they present different mortality levels, and visually the Poisson Hierarchical model (blue curve) is the one with closest distance[3] from the observed values, cases of São Luís, Adamantina and São Paulo. In addition to that, this model is the only one where we can identify variation in the mortality levels. In general, the models do not fit infant and child mortality very well. This is a relevant point, we are working on different set of adjustments that will be presented in the final version of the paper. For this reason, we also evaluate the quality of the estimates generated bythe 45q15(adult mortality). But, most important they are very useful to smooth mortality functions allowing to estimate life tables. In most cases, one cannot use observed data to perform any type of analysis.

Figure 8 – Parametric Models to Estimate Mortality Curves for Municipalities in Maranhão and São Paulo, Males, 2000

$The distances between observed and fitted curves was also estimated based on root mean square error (RMSE), which is a frequently used measure of the differences between values (sample and population v

!
!

4. Discussion

Adult mortality estimates for developing countries are less satisfactory than infant and child mortality estimates. There are two main reasons for that: there is not an equivalent to birth history to estimate child mortality from surveys and the methods used to estimate adult mortality from surveys are less robust than the ones used to estimate infant and child mortality. In result, we know much less about level and trends of adult mortality in developing countries than what we know about infant and child mortality.

The results about the evolution of completeness of death counts and adult

!

!

Bernardinelli, L. and Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk.Statistics in Medicine, 11: 983-1007.

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

!
!

Mathers,CD; Inoue, M; Rao, C; Lopez, AD. (2005). Counting the dead and what they died from: an assessment of the global status of cause of death data. Bulletin of the Word Health Organization 83(3), March: 171-9.

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