Improving mortality estimates from household deaths reported in censuses (Extended Abstract)

Gabriel Mendes Borges

Abstract

Questions on recent deaths in the household have been included in at least 70 censuses around the world. This information represents an important source of mortality estimation in developing countries and, since it is normally collected in censuses, there is no implicit sampling error and mortality estimators use numerator and denominator from the same source, unlike when administrative data are used. It can be also used to measure differentials in mortality according to different characteristics of the household, such as income. Nevertheless, this information has not been widely used, which is partially due to the limitations of the data, for instance under-reporting as a result of households with no survivors to report the death. This paper proposes a new method to adjust for this kind of bias. Simulations for Brazil, which has fairly reliable data for comparison, show that the proposed adjustment greatly improves mortality estimations, particularly at old ages.

Introduction

Estimating mortality in developing countries is a challenging task, due to the lack of reliable and complete vital registration systems. Different methods have been proposed to adjust death counts (e.g. the death distribution methods) and to estimate mortality directly using questions included in surveys and censuses. Two examples of this last approach are the sibling survival method and the deaths reported in the households (Timæus et al., 2013).

The question about the existence of a recent death in the household has been included in at least 70 censuses in different regions of the world (Appendix 1), such as Latin America (e.g. Brazil, Bolivia and Paraguay), Sub-Saharan Africa (e.g. Nigeria, South Africa, Rwanda) and Southeast Asia (e.g. Vietnam, Indonesia).

This information has some advantages, related to the fact that it is collected in a census, which normally does not have implicit sampling errors. Furthermore, numerator and denominator come from the same source and mortality inequality can be estimated according to different characteristics of the household collected in the censuses, for instance income, rural/urban classification and access to water supply and sanitation. This is also the only information collected in censuses and surveys that provides mortality estimates for all ages (Queiroz and Sawyer, 2012; UNSD, 2004).

Nevertheless, these data has not been much used to estimate mortality, although some successful applications have been performed in China (Banister and Hill, 2004), Brazil (Queiroz and Sawyer, 2012) and Sub-Saharan African countries: Cameroon (Bangha, 2010), Lesotho and Botswana (Thomas and Hill, 2007).

The lack of a more extensive use of the household deaths is partially due to the limitations of the data (Hill et al., 2005). Potential sources of bias in this information are related to the confusion about how to report deaths in the case of the disintegration of the household due to a death or when people are seen as belonging to more than one household (Dorrington et al., 2006).

More importantly, under-reporting as a result of households with no survivors to report the death is also thought to be a significant limitation. This is particularly clear in one-person households, where the death, by definition, cannot be reported. Additionally, this bias is also likely to be differential by age, being higher at older ages, in which one-person households are more prevalent (Dorrington et al., 2006). The United Nations, which has for many years included in its recommendations questions on recent household deaths (UNSD, 2008), also highlights that typically these reports underestimate the overall number of deaths because some deaths result in the disintegration of households so that household survivors, if any, may not report their occurrence (in particular, deaths of persons living alone at the time of death are unlikely to be reported). As a result of this main limitation, Queiroz and Sawyer (2012) show that household deaths reported in the 2010 Census in Brazil, for example, represent a good picture of the age structure of mortality, except for the older age groups.

Thus, the aim of this paper is to propose an adjustment method for this last kind of bias, often seen as the main source of error in mortality estimates from household deaths reported in censuses.

Method

The method proposed in this paper builds on the existing literature about corrections for selection biases in the sibling survival method, such as the resulting from the fact that some sibships are not observed (Gakidou and King, 2006; Masquelier, 2012), in addition to the recent works on the network reporting framework for estimating adult mortality (Feehan et al., 2016). The underlying idea of proposed adjustment is to reconcile the numerator (reported deaths) and the denominator (population exposed to the risk of dying).

The first adjustment, perhaps obvious, but often neglected, is to exclude from the denominator the population living in households where the question about household deaths was not asked, for instance people living in institutions.

The second adjustment intends to correct for the bias of under-reporting as a result of households with no survivors to report the death. Figure 1 illustrates this source of bias by showing the population counted in the census (P_t) by household size, the number of deaths occurred in the years before the census date ($D_{t-1,t}$) and the possible outcomes for the population at time $t-1$ (P_{t-1}) given the population and deaths reported in the census. It is worth noting that P_{t-1} is the actual population exposed to the risk of dying. The solid lines depict the quantities observed in the censuses and the non-observed ones are represented by the dashed lines, which are the cases in which all the household members died within the years before the census. The lack of this information clearly bias mortality estimation downward, since part of the population at risk is still being reported, although the deaths are not.

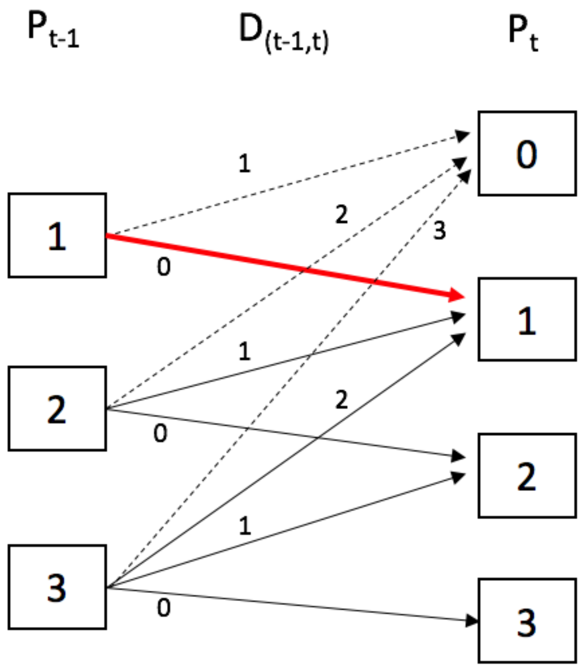

Figure 1 – Diagram showing the population by number of household members in the census date (P_t), the number of household deaths reported in the last 12 months ($D_{t-1,t}$), and the population estimated for one year before the census date (P_{t-1}).
Note: this diagram can be similarly extended for more than 3 household members.

One way of adjusting for this bias would be through the estimation of the deaths that occurred in households in which all members died. Assuming that the joint probability of dying of all household members within a year is significantly low for households with two or more members, we turn our attention to the estimation of the deaths that would have occurred only in one-person households.

Alternatively, instead of trying to estimate the deaths in households with no survivors, this paper proposes an adjustment in the denominator, which is the exposed-to-risk population. Since we cannot observe the first dashed arrow in Figure 1 ($1\rightarrow 0$), a more consistent denominator for the mortality rate would be the one which removes the population exposed to the risk, given by the red solid arrow ($1\rightarrow 1$). Note that the new population estimate P_t^* would now exclude not all on-person households in the census, but only those with no deaths reported. The main assumption of this procedure is that mortality does not vary significantly by household size.

The naïve mortality estimation for the year before the census by age and sex group α, which has been usually calculated, is given by the total number of reported deaths $D_{(t-1,t),\alpha}$ divided by the population in the same group counted in the census ($P_{t,\alpha}$):

$$M_{(t-1,t),\alpha} = \frac{D_{(t-1,t),\alpha}}{P_{t,\alpha}}$$

The adjusted mortality estimation is given by:

$$M^*_{(t-1,t),\alpha} = \frac{D_{(t-1,t),\alpha}}{P^*_{(t-1/2),\alpha}} = \frac{D_{(t-1,t),\alpha}}{P_{t,\alpha}+0.5D_{(t-1,t),\alpha}-P_{t,\alpha,(h>1,d=0)}},$$

where $P^*_{(t-1/2),\alpha}$ is the population at time *t* ($P_{t,\alpha}$), plus half of the reported deaths ($0.5D_{(t-1,t),\alpha}$), minus the population at time *t* living in one-person households with no reported deaths in the 12 months period prior to the census ($P_{t,\alpha,(h>1,d=0)}$).

To test the accuracy of this adjustment, the next section presents the results for the 2010 Census in Brazil, which has relatively reliable mortality estimations using different data sources, which are used for comparison. In addition to compare to the official mortality estimations for the country as a whole, the adjusted mortality estimations can be also compared to the mortality calculated using vital registration in states where this information is known as of good quality.

Preliminary Results

Figure 2 compares the unadjusted and adjusted mortality estimates with the official estimates published by the Brazilian Census Bureau (IBGE), which is calculated using vital registration data and adjusted for undercount of deaths using indirect demographic techniques (IBGE, 2013).

The comparison shows that mortality rates estimated using the adjustment method proposed in this paper are much closer from the official estimates than the unadjusted rates. For females, mortality rates practically coincide with official estimates for all age groups above 25 years old, whereas the unadjusted estimates diverge from the “true” value as age increases. It should be noticed that, since the official estimates correct the registered deaths, this information is also subject to errors.

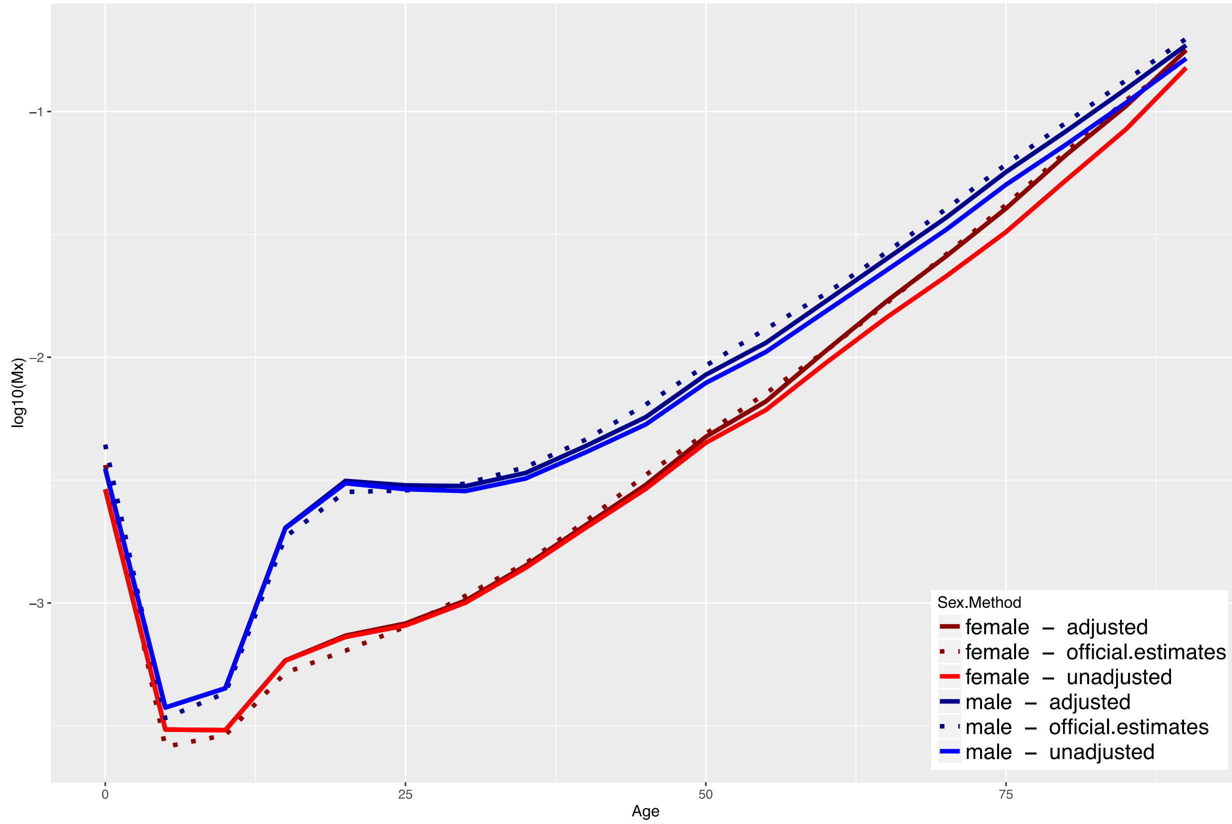

Figure 2 – Age Specific Mortality Rates by age, sex and method – Brazil, 2010

Figure 3 shows the comparison of the unadjusted and adjusted mortality estimates with the mortality rates calculated using death counts from the vital registration system for São Paulo, the most populous and one of the most developed states in Brazil, which is also known for having a good vital registration system (RIPSA, 2012). Once again, the use of the adjustment seems to improve the estimates.

It is worth mentioning that household deaths information appears to overestimate mortality among children and young adults from 5 to 25 years old.

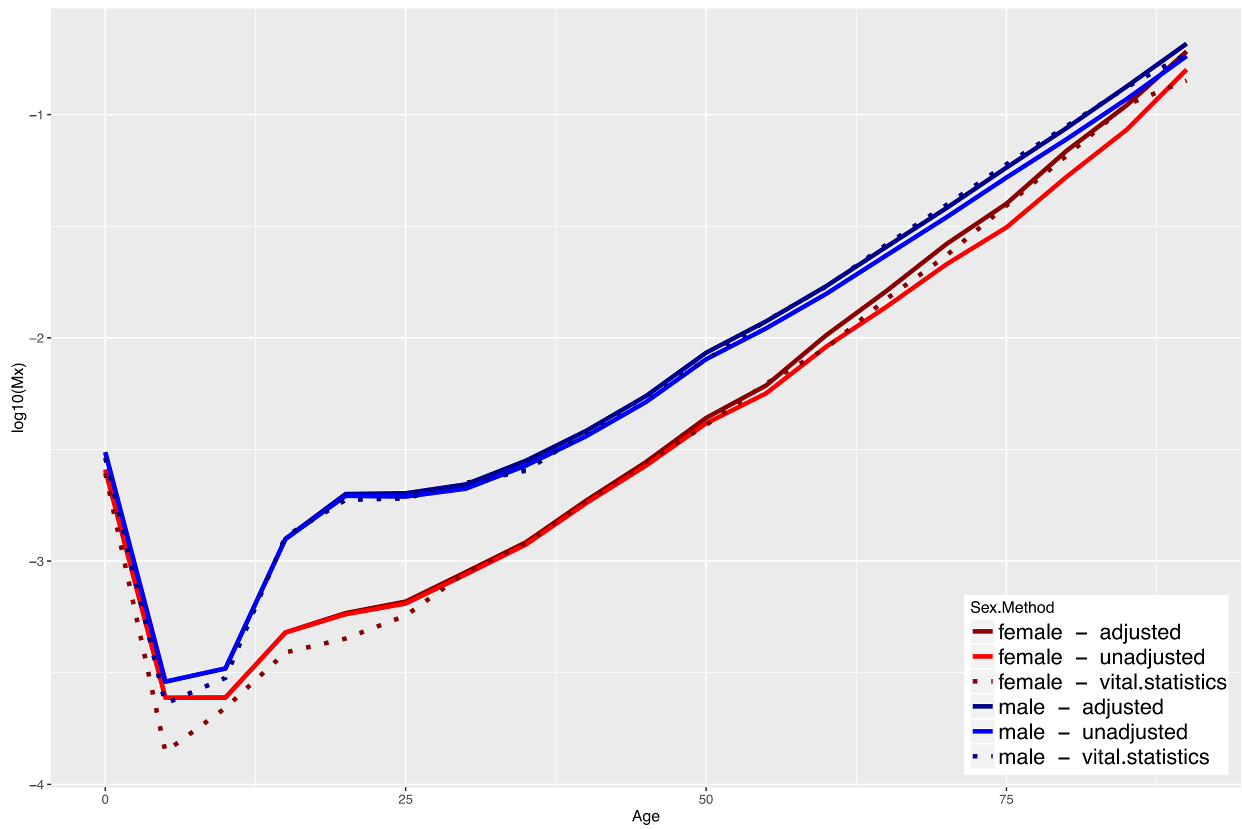

Figure 3 – Age Specific Mortality Rates by age, sex and method – São Paulo, 2010

Finally, Figure 4 depicts an important application of the deaths in the household and the method proposed in this paper, comparing mortality rates by household income. As expected, mortality rates are substantially higher for people living in poverty conditions.

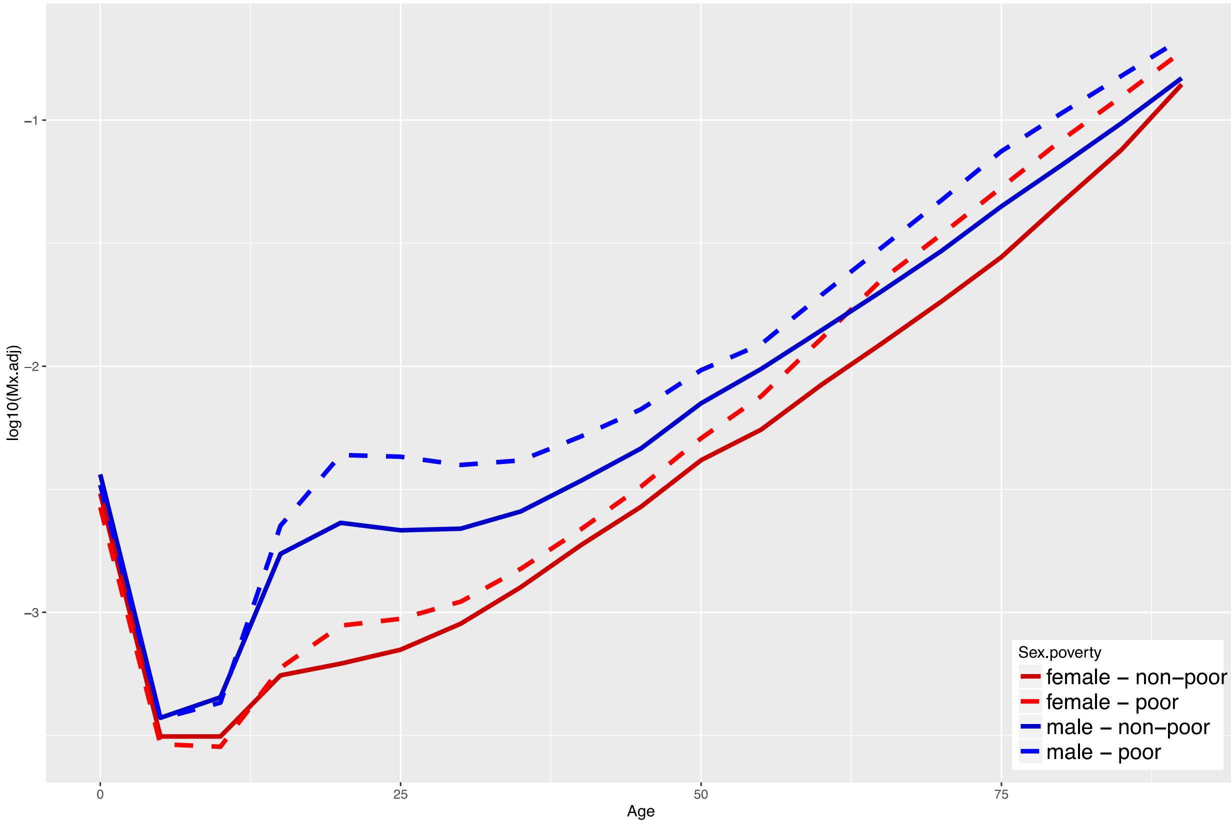

Figure 4 – Age Specific Mortality Rates by age, sex and poverty condition – Brazil, 2010

Discussion and Conclusion

This paper has proposed a method to adjust for bias on recent deaths in the household reported in censuses. Application for Brazil shows that the adjustment greatly improves mortality estimations, particularly at old ages, resolving one of the main limitations of this information.

The application of the proposed method to estimate mortality in Brazil show figures extremely consistent with the best information available, showing no need to further adjustments. For other contexts, however, adjusted mortality rates might need additional corrections, which could be accomplished by the existing methods for assessing and adjusting completeness of deaths counts from vital statistics. In this case, since these methods normally assume constant adjustment factor over age, the method proposed in this paper will still be useful as it corrects for the underreported biases by age.

This method can be easily applied to improve mortality estimations in other countries that have collected these data, such as in Sub-Saharan Africa. The final paper will extend the application of the proposed method for other countries (for instance, South Africa 2011 and Rwanda 2002), assessing the sensitivity of the adjustment and comparing the results with other existing mortality estimations.

It will also discuss other sources of error, such as over-reporting (when deaths are reported by more than one household) among children and young adults and the possible changes in questionnaire design that might avoid such a problem.

Appendix 1 – Censuses that have collected information on recent deaths in the household

Benin: 1992, 2002	Haiti: 1982, 2003	Panama: 2000
Bolivia: 1992, 2001, 2012	Honduras: 2001, 2013	Paraguay: 2002, 2012
Botswana: 1991	Indonesia: 1976, 1985, 2005, 2010	Rwanda: 2002
Brazil: 1980, 2010	Iran: 1996	Senegal: 2002
Burkina Faso: 1985, 1996	Jamaica: 2001	Sierra Leone: 2004
Cambodia: 2008	Kenya: 2009	South Africa: 2007, 2011
Cameroon: 1976, 1987, 2005	Laos: 1995	South Sudan: 2008
China: 1982, 1990, 2000	Lesotho: 1996	Sudan: 2008
Colombia: 2005	Liberia: 2008	Tanzania: 1988, 2002
Dominican Republic: 2010	Madagascar: 1993	Uganda: 2002
El Salvador: 1992, 2007	Malawi: 1998, 2008	Vietnam: 1989, 1999
Ethiopia: 1984, 2007	Mali: 1987, 1998, 2009	Zambia: 1969, 1990, 2010
Ghana: 2010	Mozambique: 2007	Zimbabwe: 1992
Guatemala: 1994	Nicaragua: 2005	
Guinea: 1996	Nigeria: 2006-2007, 2009	

Source: Minnesota Population Center (2015) and other censuses' questionnaires.

References

Bangha, M.W. (2010). Estimating adult mortality in Cameroon from census data on household deaths: 1976-1987.

Banister, J., and Hill, K. (2004). Mortality in China 1964-2000. Popul Stud (Camb) *58*, 55–75.

Dorrington, R.E., Timaeus, I.M., and Gregson, S. (2006). Adult mortality in Southern Africa using deaths reported by households: Some methodological issues and results. In Annual Conference of the Population Association of America. Los Angeles, p.

Feehan, D.M., Mahy, M., and Salganik, M.J. (2016). The network survival method for estimating adult mortality: Evidence from a survey experiment in Rwanda. arXiv Preprint arXiv:1603.08588.

Gakidou, E., and King, G. (2006). Death by survey: estimating adult mortality without selection bias from sibling survival data. Demography *43*, 569–585.

Hill, K., Choi, Y., and Timæus, I.M. (2005). Unconventional approaches to mortality estimation. Demographic Research *13*, 281–299.

IBGE (2013). Projeções da População: Brasil e Unidades da Federação (Rio de Janeiro: IBGE).

Masquelier, B. (2012). Adult mortality from sibling survival data: a reappraisal of selection biases. Demography 1–22.

Minnesota Population Center (2015). Integrated Public Use Microdata Series, International: Version 6.4 [Machine-readable database]. University of Minnesota, Minneapolis. OpenURL.

Queiroz, B.L., and Sawyer, D.O.T. (2012). What can the mortality data from the 2010 Census tell us? Revista Brasileira de Estudos de População *29*, 225–238.

RIPSA (2012). Indicadores e Dados Básicos - Brasil - 2012.

Thomas, K., and Hill, K. (2007). What can data on household deaths tell us about adult mortality in Lesotho and Botswana? (Arusha, Tanzania), p.

Timæus, I.M., Dorrington, R., and Hill, K. (2013). Introduction to adult mortality analysis. In Tools for Demographic Estimation, (Paris: International Union for the Scientific Study of Population), p.

UNSD (2004). Handbook on the collection of fertility and mortality data (United Nations Publications).

UNSD (2008). Principles and recommendations for population and housing censuses (United Nations Publications).