Marília Regina Nepomuceno

Overcoming the Limitations of Demographic Data: Papers on Mortality, the Extreme Aged and Education

Belo Horizonte
Federal University of Minas Gerais
Center of Development and Regional Planning
2017

Marília Regina Nepomuceno

Overcoming the Limitations of Demographic Data: Papers on Mortality, Extreme Aged Populations and Education

A thesis submitted in partial fulfillment for the doctoral degree in Demography in the Center for Development and Regional Planning of the Federal University of Minas Gerais.

Supervisor: Prof. PhD. Cássio Maldonado Turra

Belo Horizonte
Federal University of Minas Gerais
Center of Development and Regional Planning
Graduate Program in Demography
2017

Ficha catalográfica

N441o
2017

Nepomuceno, Marília Regina.
Overcoming the Limitations of Demographic Data: Papers on Mortality, the Extreme Aged and Education [manuscrito] /Marília Regina Nepomuceno . – 2017.
69 f.: il., gráfs. e tabs.

Orientador: Cássio Maldonado Turra.
Tese (doutorado) – Universidade Federal de Minas Gerais, Centro de Desenvolvimento e Planejamento Regional.
Inclui bibliografia (f. 53-68).

1. Demografia – Teses. 2. Brasil – Censo demográfico – Teses. 3. Mortalidade – Teses. 4. Idosos – Brasil – Teses. I. Turra, Cássio Maldonado. II. Universidade Federal de Minas Gerais. Centro de Desenvolvimento e Planejamento Regional. III. Título.

CDD: 304.6

Elaborado por Rosilene Santos CRB-6/2527
Biblioteca da FACE/UFMG. –RSS148/2019

Curso de Pós-Graduação em Demografia da Faculdade de Ciências Econômicas

ATA DE DEFESA DE TESE DE **MARÍLIA REGINA NEPOMUCENO MARINHO** Nº. REGISTRO 2012664533. Às nove horas e trinta minutos do dia onze do mês de setembro de dois mil e dezessete, reuniu-se na ***Faculdade de Ciências Econômicas*** da Universidade Federal de Minas Gerais a Comissão Examinadora de TESE, indicada "*ad referendum*" pelo Colegiado do Curso em 24/07/2017, para julgar, em exame final, o trabalho final intitulado **"Overcoming the Limitations of Demographic Data: Papers on Mortality, the Extreme Aged and Education"**, requisito final para a obtenção do Grau de *Doutor em Demografia,* área de concentração em Demografia. Abrindo a sessão, o Presidente da Comissão, Prof. Cássio Maldonado Turra, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata, para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após, a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição do resultado final. A Comissão aprovou a candidata por unanimidade. O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 11 de setembro de 2017.

Prof. Cássio Maldonado Turra
(Orientador) (CEDEPLAR/FACE/UFMG)

Profa. Simone Wajnman
(CEDEPLAR/FACE/UFMG)

Prof. José Alberto Magno de Carvalho
(CEDEPLAR/FACE/UFMG)

Prof. Bernardo Lanza Queiroz
(CEDEPLAR/FACE/UFMG)

Prof. Marcos Roberto Gonzaga
(Universidade Federal do Rio Grande do Norte/UFRN)

Profa. Mirian Martins Ribeiro
(Universidade Federal de Ouro Preto/UFOP)

Profa. Laura Lídia Rodríguez Wong
Coordenadora do Curso de Pós-Graduação
em Demografia

Contents

Agradecimentos

Feliz, diante de mais um ciclo que se encerra, agradeço ao Pedro, companheiro de todos os momentos, Mamãe e vovó pelo imenso amor, toda minha família e amigos pelo carinho e paciência.

Agradeço ao meu orientador e amigo Cássio Maldonado Turra, aos professores do Cedeplar, em especial ao Professor José Alberto Magno de Carvalho e às Professoras Laura Lídia Rodríguez Wong e Simone Wajnman pelos ensinamentos e por me acolherem de forma tão carinhosa no departamento. Agradeço, também, a todos que, direta ou indiretamente, contribuíram com meu aprendizado ao longo da minha trajetória no Cedeplar.

Agradeço à Deus pelo bom pensamento que me inspirou.

Abstract

One of the basic tasks in demography is to produce reliable estimates. To do that, demographers require high data quality or being able to detect and correct errors. Unfortunately, in most countries data are deficient and is it necessary to be critical about the reliability of the available information. Therefore, this dissertation deals with data quality issues that are related, directly or indirectly, with survival analysis. We divided this study into three articles. The first one focuses on cohort mortality analysis in the absence of complete cohort mortality data, using the *Truncated Cross-sectional Average Length of Life* (*TCAL*). The second study examines data quality issues among the oldest-old population, focusing on the estimation of centenarians. The third article evaluates education reporting among adults in order to offer evidences of education misstatement. All the articles provide alternative ways to minimize the lack of complete cohort mortality data or the existence of inaccurate data.

Resumo

Para os demógrafos, é de grande relevância a produção de estimativas confiáveis. No entanto, para alcançar tal objetivo, é necessário dispor de dados de boa qualidade ou detectar e corrigir erros nos dados. Infelizmente, na maioria dos países, os dados contêm erros, logo, é necessário ser crítico quanto à confiabilidade das informações disponíveis. Por isso, esta tese aborda problemas na qualidade dos dados que estão relacionados, direta ou indiretamente, à análise de sobrevivência. Nós dividimos este estudo em três artigos. O primeiro centra-se na análise da mortalidade por coorte, mesmo quando não se dispõe de informações completas de mortalidade por coorte utilizando *The Truncated Cross-sectional Average Length of Life* (*TCAL*). O segundo estudo examina questões relacionadas à qualidade de dados no segmento populacional mais envelhecido, estimando, de maneira indireta, a população centenária. O terceiro artigo avalia a qualidade da educação reportada nos censos, com o objetivo de identificar erros nesta informação. Todos os três artigos que compõem esta tese contribuem com a discussão sobre a qualidade dos dados, apontando alternativas para contornar essas limitações.

1. Introduction

Demographic analysis depends on data. In places where data are reliable, consistent, and available, demographers can achieve a deeper understanding of the population dynamics. Unfortunately, in most countries data are deficient and it is necessary to be critical about the reliability of the available information. However, the lack of data or the existence of inaccurate data have never precluded demographers from finding alternative methodologies to overcome the existing limitations and to provide the best estimates possible.

This PhD dissertation deals with data quality issues that are related, directly or indirectly, with survival analysis. For computing mortality estimates, data are drawn mainly from the vital registration system and the census data. However, in many places such as Brazil, death statistics and census data are missing or have content errors. Therefore, the measurement of mortality levels and patterns can be a challenge, and may reduce our knowledge about other demographic dimensions that are related to mortality including the size of extreme-old populations and differences in survival by socioeconomic status.

This dissertation is divided into three articles. The first one focuses on cohort mortality analysis in the absence of complete cohort mortality data starting at birth, using the *Truncated Cross-sectional Average Length of Life* (*TCAL*). The second study examines data quality issues among the oldest-old population, focusing on the estimation of centenarians in Brazil. The third article evaluates education reporting among Brazilian adults in order to offer evidences of education misstatement.

The first article has been accepted to publication at the Journal *Population* of the of the *Institut National D'études Démographiques* (INED), forthcoming in the volume 74, in 2017.This article is a result of an important part of my PhD, when, for a year, I was enrolled at the European Doctoral School of Demography (EDSD). During this period, I worked with Vladimir Canudas-Romo, an associate Professor at Max-Planck Odense Center, who supervised me during this study. In this article we compare mortality of cohorts not limited to those populations that have complete mortality data starting at birth. The aim of this study is to decompose the mortality gap by age and cohort between Eastern-European countries and a group of today’s high-mortality countries (HLCs).

The second paper aims to estimate the number of centenarians in Brazil. As a result of the demographic transition, particularly due to mortality reductions from age 80 to 100, there is increasing attention devoted to the proliferation of centenarians in the world. In Brazil, the main source of information about the centenarian population is the census data. However, due to data quality issues, Brazilian demographers have been skeptical about the true number of centenarians that have been enumerated in the country. In order to evaluate the quality of census population at oldest age, we apply the variable-*r* relations to provide alternative estimates of the population of centenarians. This article was presented at the 2017 annual meeting of Population Association of America in the regular session "Global Perspectives on Health and Mortality".

Finally, the third paper aims to evaluate the quality of education census data at adult ages. To the extent that education is an important variable in demographic analysis, it is essential to know how accurate this information is. Several studies have already documented errors in education reporting, however, little is known about them in Brazil. Therefore, the aim of this article is to offer evidences of education misstatement by analyzing three different indicators: the completeness of education data, education misstatement, and educational levels by sex. Next we calculate the survivorship ratios in order to provide further indirect evidences of education misstatement.

This dissertation is organized as follows. Chapter 2 presents the paper entitled "Cohort Survival Comparisons in Eastern-European Countries: The Truncated Cross-sectional Average Length of Life Approach." The paper "The population of centenarians in Brazil: Indirect estimates using an alternative approach" we show in Chapter 3. The third paper "Assessing the quality of self-reported education among adults in Brazil, 1991-2000" we present in Chapter 4. The Final Remarks are presented in Chapter 5.

2. A Cohort Survival Comparison between Central–Eastern European and High-Longevity Countries

2.1. Introduction

Improvements in mortality do not take place uniformly across regions. Generally, mortality is higher in places where the standard of living is lower. In Europe, it is well-documented that the mortality levels of the Central and Eastern European countries are far higher than those of more developed regions (WHO, 1995; Mustard, 1996; Meslé, 1996; Velkova et al. 1997; Meslé and Vallin, 2002; Andreev et al. 2003; Meslé, 2004). For instance, in the early1990s, the mortality gap stood at over 10 years between Eastern European countries with the lowest life expectancies and Western ones with the highest (Bobak and Marmot, 1996). Yet, despite the great survival improvements in the Eastern European countries, mortality differences persist in Europe (Leon, 2011; Shkolnikov et al. 2013; Mackenbach, 2013; Meslé and Vallin, 2017).

The mortality gap between Central–Eastern and Western Europe is due mostly to changes in health and disease patterns over time (Bobak and Marmot, 1996; Andreev et al. 2003; Meslé et al. 2012; Nolte et al. 2000), and it is usually attributed to differences in socioeconomic, environmental, and public-health investments (Watson, 1995; Bobak and Marmot, 1996; Bobak, 1996; Forster, 1996; Velkova et al. 1997; Vogt et al. 2017). From the end of World War II to the mid-1960s, the increasing use of antibiotics and immunization led to Central and Eastern Europe (CEE) achieving huge progress in survival from infectious diseases, particularly among the youngest ages (Meslé and Vallin, 2002; Vallin and Meslé, 2004). At this time, CEE countries converged towards lower mortality levels, and some almost succeeded in catching up with Northern and Western European countries (Meslé and Vallin 2002; Meslé 2004). Then, in places where mortality at the youngest ages had already reached low levels, a new challenge emerged in the form of improving longevity (Vallin and Meslé, 2004; Canudas-Romo, 2010; Bergeron-Boucher et al. 2015). Following the epidemiologic transition (Omran, 1971), more developed regions showed great progress in survival by degenerative and man-made diseases. Conversely, in the CEE countries, particularly in the former USSR, where high mortality hit adults the hardest, from the mid-1960s to the mid-1980s, adult and old mortality increased for men and stagnated for women (Bobak and Marmot, 1996; Shkolnikov et al. 1997; Meslé and Vallin, 2002). As a result, mortality began to diverge between Central–Eastern and Western Europe (Vallin and Meslé, 2004; Shkolnikov, 2004).

From the late 1980s, a new divergence in mortality trends began to emerge within the CEE countries, thus producing a clear gap between Central and Eastern Europe. In Central Europe, health improvements reduced mortality from cardiovascular disease, which in turn increased life expectancy (Meslé 2004). In contrast, the former USSR countries, followed a brief period of improvement (1985–1986), with a sharp increase in mortality due to the economic crises of the 1990s and low investments in public health (Leon et al. 1997; Meslé, 2004; Andreev et al. 2003; Shkolnikov, 2004; Gavrilova et al. 2001; Grigoriev et al. 2010).

However, since the beginning of the 21st century, a new mortality trend has been observed in the former USSR countries, where mortality from cardiovascular disease and external causes of death at adult ages have started to decline (Grigoriev et al. 2010; Jasilionis et al. 2011; Shkolnikov et al. 2013; Grigoriev et al. 2014, Grigoriev and Andreev, 2015). In Russia, for instance, life expectancy at birth for both sexes increased by more than 5 years between 2004 and 2014, reaching 70.91 years in 2014,– the highest level in the country's recent history (Human Mortality Database, 2017). Another huge survival improvement took place in Belarus, where male life expectancy increased by about 2 years in one calendar year (2011–2012) (Grigoriev and Andreev 2015). Despite this recent great mortality improvement in the former USSR countries, in terms of health, their populations still lag behind Western European countries.

As already mentioned, more developed countries, like those in Western Europe, have shown long-term improvements in health, while also experiencing a sustainable decline in mortality over the past decades. From a cohort perspective, a continuous decrease in mortality over time leads to several generations experiencing the gradual benefits of health improvements, with younger cohorts benefiting more because their mortality experience begins at levels lower than those of older cohorts.

In the case of discontinuous or short-term health progress, on the other hand, only a few cohorts can enjoy the benefits of health advances. The anti-alcohol campaign that the Soviet Union established in a specific time frame (1985–1986) provides an example of this. The campaign reduced the mortality of several cohorts at adult ages during the mid-1980s, leading to a short period of increasing life expectancy (Shkolnikov and Nemtsov, 1997; Shkolnikov et al., 2004; Shkolnikov, 2012; Grigoriev and Andreev, 2015). Then, the economic crises of the 1990s hit certain age groups in the former USSR countries harder than others, with some cohorts experiencing greater negative health effects than others (Shkolnikov, 2012).

German reunification also exemplifies the importance of period changes in reducing mortality. After the fall of the Berlin Wall, mortality improved for all age groups in East Germany, but at a differential pace among ages (Vogt, 2013; Vogt and Kluge, 2015, Vogt et al. 2017), which suggests that the mortality gap differs across birth cohorts. For instance, the Eastern European cohorts born in 1890, 1900 and 1910 converged to Western European mortality levels faster than the younger cohorts (born in 1920 and 1930) (Vogt and Missov, 2017). Thus, we can speculate that the mortality differences between East and West German older cohorts fell faster than those between younger cohorts. This indicates that the contributions to mortality gap during the East–West Germany differ across birth cohorts in Germany.

Given the different political regimes, the fact that Central–Eastern European socioeconomic and medical policies have differed from those of Western Europe over recent decades, and because changes have had major effects on some CEE birth cohorts but not on others, we wondered whether the mortality gap between CEE and Western European countries also varies across birth cohorts. Further, we hypothesized that the contribution of each birth cohort to the overall mortality gap in a given time period can vary.

Our aim here is to compare mortality between CEE and a group of high-longevity countries (HLC) through a measure similar to period life expectancy, but based on available cohort survival data. To accomplish this, we calculate and then decompose the Truncated Cross-sectional Average Length of Life (TCAL) measure (Canudas-Romo and Guillot, 2015). By decomposing the gap in TCAL by age and cohort, we can analyse the survival trajectories of CEE cohorts in comparison with their HLC counterparts. Moreover, such decomposition, makes it possible to identify the short- and long-lasting effects of survival advantage/disadvantage at a given age for a certain CEE cohort. We complement previous studies by adding cohort mortality dynamics in order to help form a better understanding of the mortality gap between CEE and HLCs.

2.2. Methods

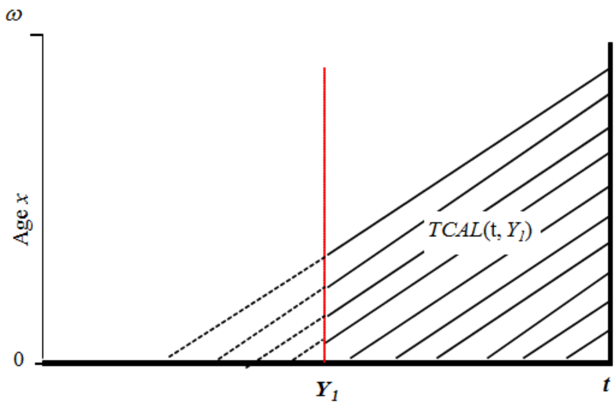
The *TCAL* is a cross-sectional measure that summarizes historical mortality information about all cohorts present at a given time, and it is not limited to populations with complete cohort mortality data (Canudas-Romo and Guillot, 2015). It derives from the *Cross-sectional Average Length of Life* (*CAL*) measure, developed by Brouard in 1986 (Brouard, 1986). *CAL* can be interpreted as the mean length of life lived by an average cohort present in a given period, in terms of the population's mortality experience (Guillot, 2003). The difference between *CAL* and *TCAL* is that the second one can be calculated for populations without complete cohort mortality data. *TCAL* provides a novel way of comparing mortality and investigating survival disparities between populations by considering all the information availabale for all cohorts presente at a given time–regardless of whether or not they have complete cohort data, and regardless of whether the data come from a young or old cohort. The *TCAL* measure has an advantage over assessments of just a single year, which, from a period analysis perspective, combines pieces of mortality information from different cohorts. It also has an advantage over assessing, one by one, each individual cohort present at a given time while not knowing how they jointly contribute to the overall survival disparities between populations.

To calculate the $TCAL$, we define the year, *t*, for which we are interested in estimating the measure, and also the earliest year for the available mortality series, Y_1. Thus, the *TCAL* for year *t*, truncated at year Y_1, is computed as:

$$TCAL(t,Y_1) = \int_0^{\omega} \ell(x,t,Y_1)dx \qquad (1)$$

where $\ell(x,t,Y_1)$ is the survival function for cohorts reaching age *x* in year *t*, whose members were born in year $t-x$. In the Lexis diagram shown in Figure 1, we see that $TCAL(t,Y_1)$ includes mortality rates from year Y_1, which are located along diagonals that cross the age axis at time *t*.

Note that some of the cohorts were born after year Y_1 and have full cohort information. For cohorts born before the year Y_1, only partial cohort mortality data are available; so we assume a set of death rates for the years before year Y_1. Since our interest is in the mortality gap between populations, the *TCAL* differences will be consistent if we use the same set of death rates for the years before Y_1 in all the examined countries (Canudas-Romo and Guillot, 2015). In order to eliminate any confounding effects of death rates before the year Y_1, we assume death rates equal to zero for all the years before Y_1, thereby focusing our comparisons solely on the cohort information available.



Source: Author's illustration.

Figure 1 – Lexis diagram for the location of death rates used in $TCAL(t,Y_1)$

To compare two populations at time *t*, both *TCALs* must be truncated at the same year (Y_1), which means, in this case, that the mortality series for all CEE countries and for the group of HLCs must start atY_1. Thus, after comparing the *TCALs* of each CEE country with the group of HLCs, we see which populations had higher mortality levels according to historical mortality data. Lower *TCAL* values correspond to populations that had higher cohort mortality levels.

The difference in *TCALs* between the group of HLCs and each CEE country, *i*, is then:

$$TCAL_{HLC}(t,Y_1)-TCAL_i(t,Y_1)=\int_0^{\omega}\left[\ell_{HLC}\left(x,t,Y_1\right)-\ell_i\left(x,t,Y_1\right)\right]dx \qquad (2)$$

where the integral corresponds to the cohorts present at time *t*, aged 0 to ω, and both populations have the same set of age-specific death rates in year Y_1. The cohort survival differences on the right side of equation (2) allow us to identify the mortality contribution of each cohort present in year *t*. The difference between *TCALs* is comparable to the difference between life expectancies in that it shows the number of years one population lags behind another.

We can rewrite equation (2) using the definition of cohort survival as:

$$TCAL_{HLCs}(t,Y_1) - TCAL_i(t,Y_1) = \int_0^{\omega} e^{-\int_0^x \mu_{HLC}(a,t-x+a)da} - e^{-\int_0^x \mu_i(a,t-x+a)da} dx \quad (3)$$

where $\mu_{HLCs}\left(a,t-x+a\right)$ and $\mu_i\left(a,t-x+a\right)$ are the forces of mortality at age *a* and time *t-x+a* for, respectively, the HLCs and population *i*. As the *TCAL* condenses the available cohort mortality history into one measure, equations (2) and (3) show that any differences between *TCAL*s allow us to identify cohort-specific contributions to the mortality gap. Thus, the age-cohort contribution $\Delta\left(a,t-x,i\right)$ to the difference between the $TCAL_{HLCs}$ of the HLCs and that of the population *i*, $TCAL_i$ can be estimated as:

$$\Delta(a,t-x,i) = \left[\frac{\ell(x,t,Y_1,HLCs) + \ell(x,t,Y_1,i)}{2}\right] \ln\left[\frac{{}_1p_a(t-x,HLCs)}{{}_1p_a(t-x,i)}\right], \quad (4)$$

where $\ell(x,t,Y_1,i)$ and $\ell(x,t,Y_1,HLCs)$ are the survival functions for the cohort aged *x* at time *t* in, respectively, the group of HLCs and population *i*; and ${}_1p_a(t-x,HLCs)$ and ${}_1p_a(t-x,i)$ are the probabilities of surviving from age a to $a+1$ for the cohort born in year $t-x$ in, respectively, the HLCs and population *i*. Finally, instead of the integrals in equations (2) and (3), the sum over cohorts and ages of the age-cohort contributions, $\Delta(a,t-x,i)$, returns the difference in *TCALs*

$$TCAL_{HLCs}(t,Y_1) - TCAL_i(t,Y_1) \approx \sum_{x=1}^{\omega}\sum_{a=0}^{x-1}\Delta(a,t-x,i)\,. \quad (5)$$

By means of such decomposition, we compare mortality between birth cohorts from different populations.

The main limitation of the method is data availability. In principal, we would be interested in presenting as much cohort data as possible. However, this is not possible for many regions of the world, such as in the CEE. Furthermore, despite constraints on data quantity, data quality improves over time. Thus, any measure with a cohort perspective will include some of the quality bias that exists in the older information.

2.3. Data

From the Human Mortality Database (2017), we selected 11 Central and Eastern European (CEE) countries: Belarus, Bulgaria, Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Russia, Slovakia, and Ukraine. The high-longevity countries (HLCs) included in the analysis are: Australia, Austria, Belgium, Canada, Denmark, Germany, Finland, France, Iceland, Ireland, Italy, Japan, the Netherlands, Norway, New Zealand, Portugal, Spain, Sweden, Switzerland, the United States, and the United Kingdom. The same selection of HLCs has been used elsewhere (Ho and Preston, 2010; Canudas-Romo and Engelman, 2012; Canudas-Romo and Guillot, 2015) to represent the lowest mortality levels. For this group, we calculated period age-specific death rates by adding annual death counts and exposures from each country.

We used mortality series from 1959 to 2013, except in the cases of Canada and Bulgaria, where HMD data are currently available only up to, respectively, 2011 and 2010. In order to compare mortality levels between each CEE country and the group of HLCs, we truncated all series in 1959, the first year of available HMD data for most of the CEE countries; then, we calculated the TCAL in 2013, truncated in 1959 – or, as expressed in equation (1), $\mathrm{TCAL}\,(2013, 1959)$.

The HMD data provide detailed historical information on mortality for most industrialized countries. It should be noted, however, that the quality of data for 1959–1969 is lower than in later years for Latvia, Russia, Ukraine and Belarus (Jasilionis, 2017; Shkolnikov and Jdanov, 2016; Pyrozhkov et al. 2015; Gregoriev, 2015). Also, data for 1959–1979 in Lithuania should be used with caution, while the quality of the data in Estonia for 2001–2009 is lower than in previous years (Jasilionis, 2017; Jasilionis, 2017). Information from earlier years (for old ages) should also be treated with caution, due to data quality issues in the former USSR countries. In Central Europe, the quality of data in Slovakia from 1959 to 1961 and in Hungary for 1959 is lower than in later years (Mészáros and Jasilionis, 2015; Jasilionis and Radnóti, 2016).

It is also relevant to mention here that the HMD publishes cohort death rates only for cohorts that have at least 30 years of data. For instance, period death rates for Ukraine are currently available up to 2013, but cohort death rates are available only up to the cohort born in 1983. Therefore, to avoid an interruption in the cohort series, we used period death rates to reconstruct the diagonals. This was done consistently across all the countries analysed, for both CEE countries and HLCs.

We further carried out a sensitivity analysis using data for countries that have enough cohort data to construct complete (i) cohort life tables, which we then compared with (ii) cohort life tables based on period rates in a diagonal manner. Based on the HMD data and combining females and males for the 1900 cohort, disparities between life expectancies at birth for each country were minor: (i) 59.3 and (ii) 59.2 years for Denmark, (i) 55.1 and (ii) 55.0 years for the Netherlands; (i) 59.6 and (ii) 59.4 years for Norway; and (i) 58.8 and (ii) 58.7 years for Sweden. Both the sensitivity analysis and the consistency of the procedure for all countries reassured us that our results are not biased by the set of death rates selected.

2.4. Results

Table 1 gives rankings of life expectancy at birth from highest to lowest. The table covers life expectancy at birth (e0), based on current mortality in 2013, and TCAL (2013, 1959), which captures the available mortality history from 1959 of all cohorts present in 2013, for each CEE country by sex. As expected, all CEE countries lag behind the group of HLCs. In comparing e0 with TCAL, Table 1 displays TCAL values that are lower than e0 for all countries. This is explained by TCAL taking into account the higher mortality levels in the historical mortality data while they are not considered in e0.

When the Central and Eastern European countries are ranked according to TCAL and e0, the Central European countries are clustered at the top, while the Eastern European countries are at the bottom. This suggests that both historical and current mortality in Eastern Europe are higher than those in Central Europe. For instance, Ukraine and Russia have the highest mortality levels according to both the TCAL and e0, while the Czechia and Poland are in the top 3 for both rankings. Despite this overall picture, we also identify some country-specific arrangements in Table 1. For instance, Estonian women move from the top of the life expectancy ranking (e0 = 81.3 years), to 3rd position in the TCAL (77.9 years); Estonian men also go down in the rankings when transiting from e0 to the TCAL, from 4th (e0 = 72.7 years) to 6th (TCAL = 66.7

years). Since the beginning of the 21st century, Estonia has shown remarkable mortality improvements. From 2000 to 2010, Estonian life expectancy increased by 4.2 years for women and 5.2 years for men. Indeed, these recent increases in life expectancy were three times and 11 times higher than over the previous decade for, respectively, women and men. This exemplifies the existing differential between current and cohort mortality, as depicted by life expectancy and TCAL.

In order to show the performance of CEE countries in relation to the HLCs, Table 1 presents the differences in e0 and in TCALs between each CEE country and the group of HLCs in 2013. In both cases, the greatest differences are between the former USSR countries and the HLCs. For instance, the Russian male TCAL was 14.1 years lower than the TCAL in HLCs in 2013, while the gap in e0 is 13.2 years. In the Czechia, however, the male gap in TCAL (Czechia vs. HLCs) is 3.8 years, about 10 years lower than the male TCAL difference between Russia and the HLCs. Considering this smaller difference in TCALs between the Czechia and the HLCs compared with the gap between Russia and the HLCs, one may conclude that the historical improvements in male survival have been much greater in the Czechia than in Russia.

Note that the male gap in life expectancy (CEE vs. HLCs) for most CEE countries is smaller than the difference in TCALs (Table 1). In Hungary, the male gap in e0 is 29% lower than the gap in TCALs, while in Estonia the male difference in life expectancy is almost 60% lower than the difference in TCALs (Estonia vs. HLCs). In other words, the difference is higher in male historical mortality than in current mortality when comparing CEE countries with HLCs. The differences in current mortality may possibly have been reduced by recent mortality improvements in the male mortality of CEE countries compared with HLCs. Among women, by contrast, Table 1 shows higher gaps in e0 (CEE vs. HLCs) than differences in TCALs (CEE vs. HLCs) for all countries except the Czechia. For instance, Lithuania’s female life expectancy is lagging behind the HLCs by 5.1 years, while the difference in cohort mortality is 20% lower. In the Ukraine, the female gap in e0 (Ukraine vs. HLCs) is more than one year higher than the difference in TCALs (Ukraine vs. HLCs). This may be explained by the slow progress over recent decades in the female mortality of CEE countries when compared with HLCs. The greater disparity in female life expectancies may possibly be explained by some past mortality improvements among women that were captured by the TCAL but not by life expectancy.

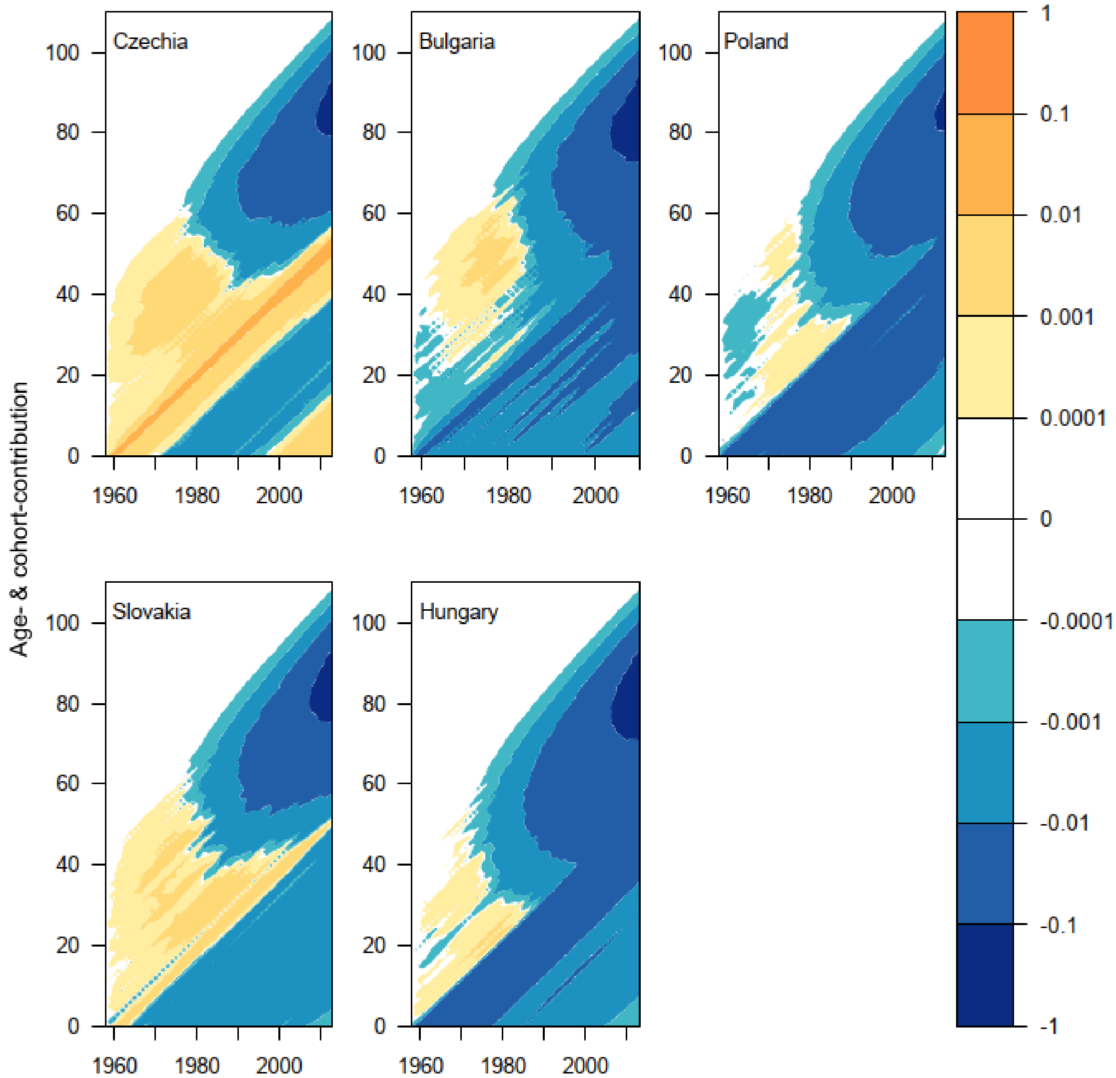
To further understand the differences in cohort survival, Figures 1 (Central European countries) and 2 (Eastern European countries) show the age-cohort decomposition of the TCAL difference between each CEE country and the group of HLCs, with females and males being indicated by, respectively, figures A and B. Such decomposition allows us to investigate the contribution that cohorts present in 2013 make to the gap in TCALs between each CEE country and the HLCs. Figures 1 and 2 show the Lexis surfaces of the cumulative age and cohort contributions to the difference in TCALs. Each data point (age-x and time-t) in these figures represents the cumulative difference in cohort survival up to the specific age-x and year-t. Negative values are associated with higher survival in the HLCs.

Figures 1A and 1B for females and males, respectively, show lower mortality levels for most cohorts in the HLCs than in their counterparts in Central Europe for both sexes. With the exception of some Czech and Polish cohorts, all Central European cohorts present in 2013 contribute to the overall mortality disadvantage between each Central European country and the HLCs in 2013.

In comparison, the overall picture of the Czechia vs. the HLCs stands out from those of the other Central European countries. Figures 1A and 2A show that Czech cohorts born in the late1950s and during the 1960s – especially those born in the early1960s –had a particular survival advantage over their counterparts in HLCs from birth until the age they reached in 2013. After World War II and up to the mid-1960s, mortality in the Czechia greatly decreased due to their extending health coverage to the entire population. At this time, Czech life expectancy at birth increased at the same rate as in France, and both countries achieved a similar mortality level (Rychtaříková, 2004). In addition to confirming the great mortality improvement during the 1960s, our decomposition reveals the long-lasting effect of lower mortality at younger ages for Czech cohorts compared with the HLCs. The low mortality in infancy and childhood lasts until 2013, and it is seen in the Czechia's higher cohort survival for cohorts born from 1959 to the early 1970s. These figures also suggest a more recent cohort development, i.e. lower infant/child mortality in the Czechia than in the HLCs. These recent mortality improvements in the Czechia may contribute to the slightly lower gap in e0 (Czechia vs. HLCs) than in TCALs.

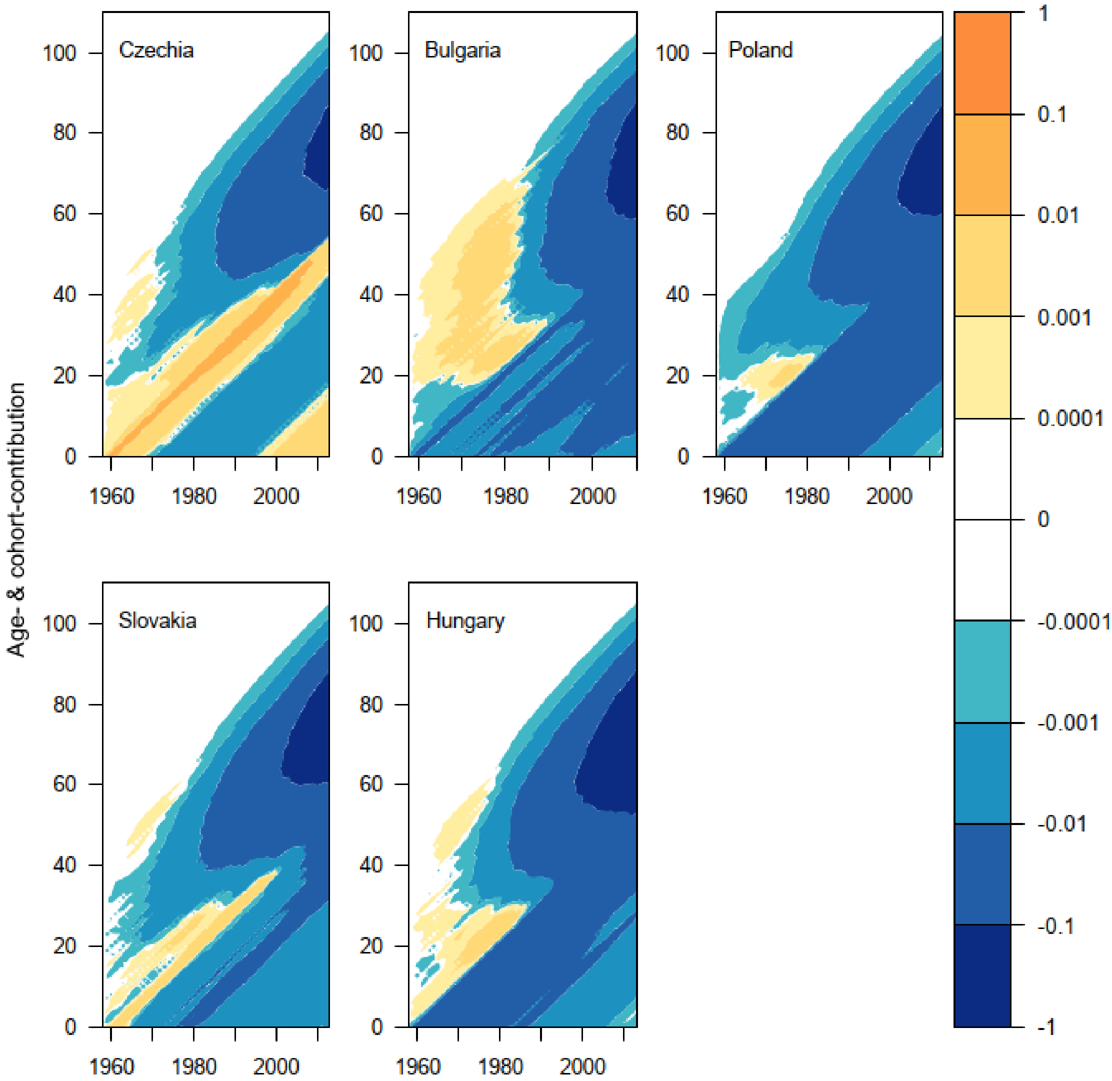
In our window of observation, from 1959 to 2013, Figures 1A and 1B also reveal that Central European cohorts born from the 1920s to the late 1950s experienced lower mortality compared with the group of HLCs. This survival advantage is greater for females than for males. However, all these Central European cohorts have gradually lost their survival advantage compared with HLCs. Figure 1B shows that the male cohorts aged 60–80 in the 2000s are the ones that contribute the most to the mortality gap between Central European countries and HLCs in 2013. Among women, despite the great contribution that the cohorts aged 60–80 in the 2000s make to the differences in TCALs, their contribution is lower in comparison to that of men.

Figure 1A - Lexis surface for the cumulative age- & cohort-contributions to the difference in TCALs between CEE countries and other HLCs, Females.



Note:Negative values correspond to higher HLCs survival.
Source: HMD data and authors'own calculation.

Figure 1B - Lexis surface for the cumulative age- & cohort-contributions to the difference in TCALs between CEE countries and other HLCs, Males.



Note:Negative values correspond to higher HLCs survival.
Source: HMD data and authors'own calculation.

For the former USSR countries (Figures 2A and 2B), the overall picture is very similar: a great survival disadvantage exists for most cohorts. Moreover, survival disadvantage (Eastern European countries vs. HLCs) for all cohorts born between 1959 and 2000 increases as the cohorts get older. Note that no Eastern European cohort has experienced lower mortality than HLCs in middle-aged adults. Indeed, the survival disadvantage of Eastern European cohorts greatly increases after age 30.

With the exception of Russia, all Eastern European cohorts born during the 1960s and the early1970sexperience a small survival advantage at younger ages than do the HLC cohorts. However, the survival advantage of these cohorts gradually disappears up to 2013 for both sexes. We also observe a longer-lasting effect of this survival advantage in infancy and childhood for females than for males. Note that male cohorts experience survival advantages at younger ages (i.e. up to age 20–25) when compared with HLCs, while female cohorts had this survival advantage until age 30-35. The low mortality levels until adulthood of Eastern European cohorts compared with HLCs was probably triggered by the anti-alcohol campaign oriented toward adults in 1985–1987, when cohorts born in the early1960shad reached ages 20–25. Moreover, our results suggest different effects on males and females as a result of the campaign, which reduced the mortality gap between each former USSR country and the group of HLCs.

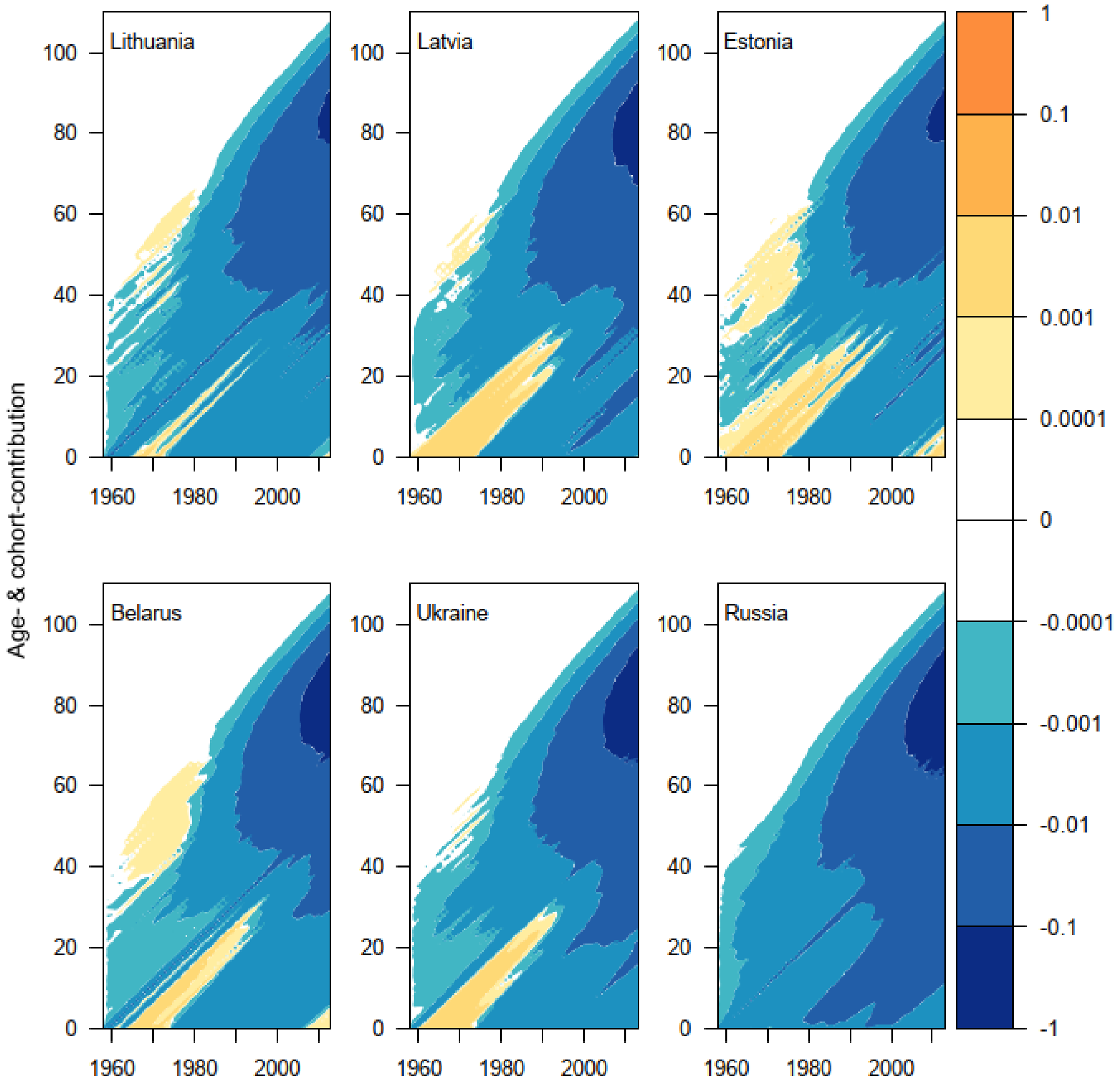
From a cohort perspective, our results also reveal that the effect of the anti-alcohol campaign differs in the former USSR countries when they compared with HLCs. In Lithuania, for instance, the survival advantage at younger ages of male cohorts born from the mid-1960s to the early1970s did not last until age 20 (Figure 2B). By contrast, in Latvia, Ukraine, and Estonia, male cohorts that experienced lower mortality in infant and childhood than HLCs retained their advantage up to ages 20–25. These aspects may indicate that – compared to changes in the mortality gap between Lithuania and the HLCs – cohort mortality improvements due to the anti-alcohol campaign more greatly narrowed the gaps between Latvia and the HLCs, Ukraine and the HLCs, and Estonia and the HLCs.

Since the 21st century began, recent cohorts are experiencing infant and child survival disadvantage in Eastern European countries when compared with HLCs. Child mortality differences in Belarus and Estonia (Belarus vs. HLCs, and Estonia vs. HLCs) narrowed rapidly for cohorts born from the late 2000s, while child survival progressed slowly in Lithuania, Russia, and Ukraine. Another relevant aspect displayed in Figure 2B for Eastern European countries, is the great contribution of male cohorts aged 40–80 in 2013 to the difference in TCALs (Eastern-Europe vs. HLCs). This result points out the important contribution of mortality over age 40 to the mortality gap between Eastern European countries and more developed countries.

Particular attention is paid to Russia, the only country that has not experienced any cohort survival advantage compared with HLCs. In addition to showing the high cohort mortality levels in Russia, Figures 2A and 2B reveal differences in the survival trajectories of Russian cohorts when compared with HLCs. At younger ages, Russian cohorts born from 1960 to the mid-1970s have experienced lower survival disadvantage than those Russian cohorts born between 1980 and 2000. The mortality disadvantage of those Russian cohorts born between 1960 to the mid-1970s lasts until age 20 for males and up to age 30 for females. As already mentioned, the anti-alcohol campaign launched by Gorbachev in 1985–1987 may explain this long-lasting effect of lower mortality difference up to adulthood between Russia and HLCs. Note that, among the former USSR countries, only Russian cohorts born during the 1960s have not experienced survival advantage when compared with HLCs. Even if the positive effect during the period of the anti-alcohol campaign was greater in Russia and led to more important mortality improvements there than in other countries, it was not enough to compensate for lower survival in the former years.

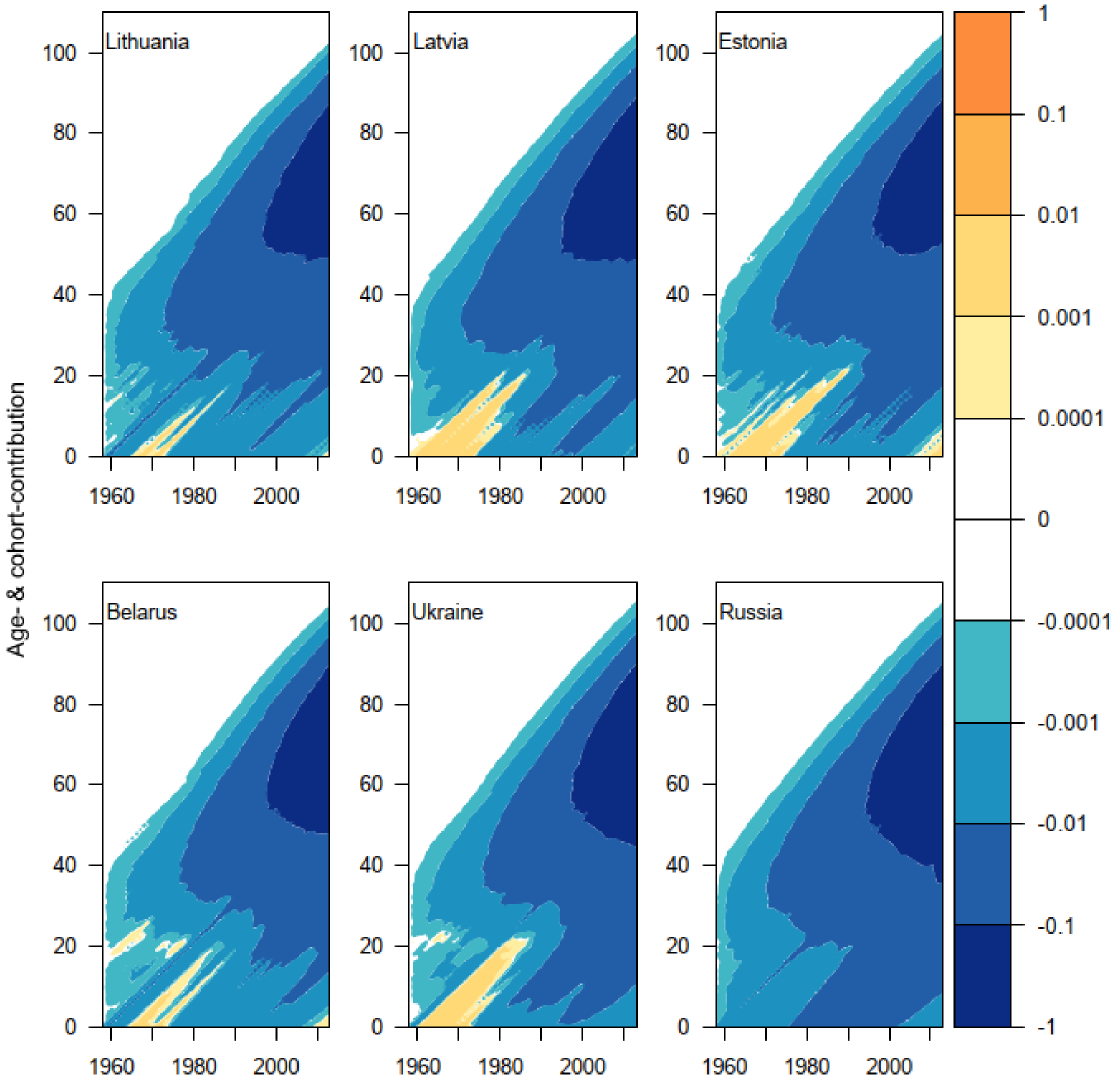
It should also be mentioned that infant and child mortality is much higher in all Russian cohorts born between 1959 and 2013 than in the HLCs. Despite the great mortality improvements at youngest ages in Russia over recent decades, the country still lags far behind when compared with HLC infant and child mortality. However, women have more quickly progressed to lower levels of infant and child mortality than have men. When compared to HLCs, survival disadvantage is lower for females than for than males at the youngest ages of Russian cohorts born during the 1980s and the 1990s.

Figure 2A - Lexis surface for the cumulative age- & cohort-contributions to the difference in TCALs between CEE countries and other HLCs, Females.



Note:Negative values correspond to higher HLCs survival.
Source: HMD data and authors'own calculation.

Figure 2B - Lexis surface for the cumulative age- & cohort-contributions to the difference in TCALs between CEE countries and other HLCs, Males.



Note:Negative values correspond to higher HLCs survival.
Source: HMD data and authors'own calculation.

2.5. Conclusion

This study takes a cohort perspective to present our findings on the mortality gap between Central and Eastern European (CEE) countries and a group of high-longevity countries (HLCs). We have revealed the contribution of cohort survival to the mortality difference between each CEE country and the group of HLCs in 2013. Our decomposition shows a great survival disadvantage for most CEE cohorts present in 2013 compared with their counterparts in HLCs. The age-cohort decomposition of difference in TCALs also reveals some survival advantages of particular CEE cohorts over HLCs, as is the case for Czech cohorts born in the late1950s and during the 1960s. These Czech cohorts had a particular survival advantage over their counterparts in HLCs from birth until the age reached in 2013. The survival advantage of these Czech cohorts confirm the documented mortality decline in the Czechia during the 1960s, when Czech life expectancy at birth was very similar to that of high-mortality countries (Rychtaříková, 2004) . We complement this result by showing the long-lasting effect of survival advantage at first ages of Czech cohorts born during the 1960s when compared with HLCs. Except for Russia, the age-cohort decomposition of the difference in TCALs also reveals a particular survival advantage of the former USSR cohorts born during the 1960s and early1970s when comparing them with HLCs. Conversely, the survival advantage of these Eastern European cohorts gradually disappears by 2013 for both sexes. Our results show that the survival advantage at younger ages for these Eastern European cohorts lasted until adulthood (up to age 20 for men, and age 30 for women). This effect was probably triggered by the anti-alcohol campaign oriented toward adults and which was launched by Gorbachev in the mid-1980s (Shkolnikov and Nemtsov, 1997), when cohorts born during the 1960s reached young-adult ages.

The TCAL decomposition helping us track how cohorts contribution to longevity evolves over time and age. The cohort perspective has been emphasized here because the aim was to understand how CEE populations arrived to current mortality levels. However, the methodology of decomposing TCAL is flexible and allows studying in the period and age perspectives by focusing instead on the age-specific contributions and accumulating correspondingly across ages or periods. Although, out of the scope of the current study, analysing all the three perspectives (age, period and cohort) together under the TCAL decomposition could complement and enrich the knowledge on population’s mortality transition.

Since the 1980s, the high mortality in Eastern European countries has been largely attributed to premature deaths in the middle-aged adult population, particularly among males born in the former USSR. (Shkolnikov et al., 1997; Shkolnikov and Nemtsov, 1997; Meslé and Vallin, 2002; Meslé, 2004) . In our window of observation, from 1959 to 2013, we show that the mortality disadvantages of the middle-aged adult population compared with other HLCs always existed between the former USSR and HLCs. Our results suggest that it is not only mortality among adults that contributes to the current disadvantage gap between the former USSR and HLCs, but that mortality at first ages still contributes to this mortality difference.

To conclude, the decomposition of the TCAL differences between CEE countries and HLCs highlights the potential for public health interventions to eliminate and control avoidable mortality gaps in the future.

3. The population of centenarians in Brazil: historical estimates from 1900 to 2000

3.1. Introduction

The number of people reaching the age of 100 has called the attention of scientists, including many demographers. In 1990, nearly 90,000 centenarians were living in the world. Over the last 25 years, this number increased five times, reaching more than 410,000 people in 2015 (U.N. 2019). Not surprisingly, given the historical and regional patterns of survival gains, there is a concentration of elderly populations in wealth low-mortality regions. However, between 1990 and 2015, the number of centenarians increased by 10% more in middle-income than high-income countries, suggesting the centenarian trend may be spreading to other regions.

Vaupel and Gowan (1986) were among the first authors to foresee the growth in the number of centenarians by simulating scenarios of survival gains in the United States. A few years later, Thatcher (1992) reported an increasing number of centenarians in England and Wales, particularly after World War II. Since then, other studies have examined trends and patterns of the centenarian population in several low-mortality countries (Barbi et al. 2018; Drefahl et al. 2012; Jdanov et al. 2008; Kestenbaum and Ferguson 2005; Leeson 2017; Medford et al. 2019; Poon and Cheung 2012; Rau et al. 2008; Robine, Saito, and Jagger 2009; Robine et al. 2010; Robine and Saito 2003; Skytthe and Jeune 1995; Thatcher 2001; Wilmoth 1995; Wilmoth and Lundström 1996)

The interest in the factors that have driven the proliferation of centenarians has been increasing as well. According to Preston and Coale (1982), population growth is due to a combination of changes in births, migration, and mortality rates. In the case of the centenarian population, Vaupel and Jeune (1995) emphasized the impact of mortality decline on the growing number of centenarians. They found that the proliferation of centenarians is mainly due to improvements in survival from ages 80 to 100 in low-mortality countries. Other authors have also stressed the role of declines in late-adult mortality, and postponement of age-at-death, to determine the size of the centenarian population (Kannisto et al. 1994; Robine and Paccaud 2005; Robine and Cubaynes 2017). In Japan, where the distribution of deaths strongly shifted to older ages between 1980 and 2000 (Robine and Cubaynes 2017), the number of centenarians increased fourteen-fold, from 913 to 13,036 (Robine and Saito 2003).

In the decades to come, we can also expect a steady proliferation of centenarians in middle-income countries, where the mortality transition - from infant and early-adult to late-adult and old ages – started later but it is in progress. There have been already some reports about the multiplication of the number of centenarians in China (Wang, Zeng, Jeune, & Vaupel, 1998), Cuba (Calzadilla et al. 2013), and in other countries like Mexico, Russia and Poland (Herm, Cheung, & Poulain, 2012). In Brazil, according to the census data, there were almost twice as many centenarians in 2010 (24,236) than in 1991 (13,296) (IBGE 1991; 2010). However, confounding data errors have precluded a more in-depth examination of the determinants of this type of demographic trends in middle-income countries, where the vital registration systems are usually weaker. The main reason is the low quality of population counts at older ages due to coverage and content errors, including age misreporting (Turra 2012).

The study of the relationship between the number of centenarians recorded in the census and data quality issues is not new and has received extensive attention also in wealthier nations. For example, in the United States, estimates of the number of centenarians showed that the census population was overestimated by about three times in 1960 (Myers 1966). Also, for the U.S., some studies have used death statistics and extinct generations methods to confirm the existence of data quality issues at older ages, particularly for some population subgroups such as African Americans (Siegel and Passel 1976). The norm has been age overstatement in census data relative to deaths, although the magnitude of the exaggeration has varied over time (Elo and Preston 1994; Rosenwaike 1979; Rosenwaike and Logue 1983). Not only in the U.S. but also the Canadian census has over-counted the number of centenarians, particularly the number of semi-supercentenarians (Bourbeau & Lebel, 2000).

In the case of Brazil, seventy years ago, Mortara (1949) was already very skeptical about the quality of the population enumerated by age in the 1920 and 1940 Census, particularly at older ages. He projected backward the number of births and the probabilities of surviving for different cohorts to estimate what would have been the number of native-born Brazilians by age in both years and compared them to the census figures. Concerning the centenarians, the results were astonishing: he estimated about six and nine people, respectively, for 1920 and 1940, whereas both censuses counted more than six thousand individuals.

Decades later, the official Brazilian figures are still suspicious. A simple comparison of the prevalence of centenarians shows that in the 1991 Census, it was about 30% higher (0.94 per 10,000) than the rate found, for example, in Sweden (0.69 per 10,000) despite the 1.5 years lower life expectancy at age 70 in Brazil (HMD 2019; IBGE 2010). In 2000, the prevalence increased to 1.45 per 10,000, but reduced to 1.27 per 10,000 in 2010, eventually becoming lower than the Swedish rate (1.71 per 10,000). This apparent erratic behavior on the prevalence rates in Brazil possibly reflects improvements in age reporting across censuses and birth cohorts. Tests of consistency have confirmed the data quality issues among older Brazilians. Following the studies developed in the U.S., Gomes and Turra (2009) compared the number centenarians in the 1991 Brazilian census with extinct generation estimates based on the number of cohort deaths in the following years. The results show significant discrepancies between the two data sources: about three times more centenarians in the census than registered deaths. Unfortunately, the death files are also not immune to both coverage and content errors, and it remains unclear what it is the most likely size of the centenarian population in Brazil.

The absence of reliable estimates of the centenarian population in populous countries, which have witnessed rapid demographic transitions but lack accurate data, prevents a more comprehensive discussion on the dynamics of these longevous subgroups beyond the limited set of lowest mortality countries. It also hinders local governments from planning policies to reach the demands of a growing old age population in places that usually face stricter family and public budget constraints. It is therefore valuable to estimate what would have been the probable evolution of the Brazilian centenarian population over the decades. In the following, we use variable-r relations (Horiuchi and Preston 1988; Preston et al. 2001) to estimate the centenarian population from the combination of different mortality schedules. Our results show a range of estimates which may reasonably be expected to bracket the correct number of centenarians over 1900-2000. Moreover, we point out the differences between our set of estimates and the reported number of centenarians in the Brazilian censuses between 1900 and 2000. We hope our findings stimulate the statistical agencies in Brazil to improve even further the process of obtaining information about older populations.

3.2. Data and Methods

The Variable-*r* Method

Variable-*r* relations can be of great value in the estimation of population measures when data are missing or are of bad quality. Preston and Coale (1982) show that the number of individuals from different age groups in the same year is related in terms of survival probabilities and demographic differences, which are reflected in the age-specific population growth rates. Following the discrete formula detailed in Preston et al. (2001), we estimated the number of centenarians in Brazil by sex, for the years that correspond to the census' years 1900, 1920, 1940, 1950, 1960, 1970, 1980, 1990 and 2000:

$$ {}_nN_x(t) = {}_nN_y(t)\, e^{-\int_y^x \bar{r}(a)da}\, {}_{x-y}p_y \qquad (1) $$

where ${}_nN_x$ is the population at the older age group at the census date, ${}_nN_y$ is the population at a younger age group at the census date, $\bar{r}(a)$ is the mean age-specific growth rate over time intervals, and ${}_{x-y}p_y$ is the probability of surviving from ages y to x. We applied Eq.1 in two steps to estimate the total number of centenarians. First, we calculated the population 100-109 years old, and next, the number of individuals 110 to 119 years old. We assumed no individual survived to ages beyond 119 in Brazil during the period of analysis.

The (Lower Limit) Age y

We relate the centenarian population to the number of individuals aged y to $y+n$. Thus, as a first step, we must define age y in Eq.1. Given the relevance of improvements of the late-adult survival to the proliferation of centenarians (Robine and Cubaynes 2017; Vaupel and Jeune 1995) and because age overstatement tends to increase with age in Brazil (Agostinho 2009) and elsewhere, we estimated the centenarian population from the number of fifty-year-olds. Since census data are not free of age misreporting, we used 10-year-wide age intervals to smooth any major inconsistencies and drew the number of people in the age group 50-59 from each successive census collected in Brazil (BRASIL 1927; BRASIL 1928; IBGE 1950; IBGE 1956; IBGE 1960; IBGE 1973; IBGE 1981; IBGE 1994; IBGE 2001; IBGE 2011). Although we believe this is the best choice for the parameter in Eq.1, to mitigate arbitrariness, we also ran sensitive analyses, described below, by varying the lower age group and the corresponding population and mortality measures.

Age-specific Growth Rates

Next, we estimated the set of age-specific growth rates above age *y*, for every intercensal period. Even if the size and the age distribution of the population at ages 50 and older is biased in the census, we can apply the variable-*r* relations as far as the pattern of data errors does not change significantly over time, affecting the age-specific growth rates (Preston et al. 2001). Nevertheless, to mitigate any potential bias in the growth rates from the variation of census data quality over time, which seems to have occurred at the oldest ages, we estimated annualized age-specific growth rates for ten-year age groups up to age group 80-89. After age 89, rather than using age-specific growth rates, we replaced them with the annualized mean growth rate for ages 90 and above. We tried different open-ended age intervals in the sensitivity analyses, as we explain below.

Because we wanted to make estimates separately for each census year, rather than for intercensal periods, we used two intercensal growth rates that are centered on each census date to calculate the mean growth rates. For example, for the age-group 70-79 in 2000, we used the mean growth rate at ages 70 to 79 in 1991-2000 and 2000-2010. Also, because in the first forty years of the analysis, census observations are separated by twenty years (1900-1920 and 1920-40), we calculated annualized twenty-year age-specific growth rates. Moreover, we assumed that age-specific growth rates for the period 1900-1920 apply to 1900.

Conditional Probability of Surviving

The third parameter in Eq.1 is the conditional probability of surviving. To estimate the number of centenarians, we computed life table survivorship ratios from ages 50-59 to age groups 100-109 and 110-119. Both coverage and content data errors (Agostinho 2009; Gomes and Turra 2009; Horta 2012) preclude us from estimating unbiased mortality rates at older ages directly from deaths and census data. In Brazil, as in many other countries where data are deficient, age misreporting is probably the main reason why death rates increase slower with age than in high-quality data countries (Coale and Kisker 1986; Dechter and Preston 1991; di Lego, Turra, and Cesar 2017; Preston, Elo, and Stewart 1999; Turra 2012). Therefore, to mitigate data quality issues, we applied and compared different methodological strategies to estimate the mortality functions, including *(i)* Coale-Demeny Life Table Models for the years 1900 to 2000; and *(ii)* three different mathematical representations of mortality – Gompertz, Weibull, and Kannisto models – for the year 2000.

As a first step, we calculated the survivorship ratios from the UN version of the Coale-Demeny (CD) Model Life Tables, which extended the original mortality levels to include life expectancy at birth up to 100 years (U.N. 2017). We decided not to choose one specific model from CD (South, North, West, and East). Instead, we offer estimates based on all four regions. To select the correct mortality level for each regional pattern in every census year, we used two different and complementary parameters: (*i*) the life expectancy at birth and (*ii*) the life expectancy at age 50. The life expectancy at birth (e_0), summarizes mortality conditions for all ages, and thus, it is less affected by data quality issues at older ages. Moreover, it is probably the best measure available to capture historical mortality levels in Brazil. We used e_0 calculated by the Brazilian Census Bureau (IBGE 1981; IBGE 2017), for all years (1900 to 2010). Next, we use life expectancy at age 50 for all the years since 1940 (IBGE 2000; IBGE 2010; IBGE 2017). It is an index of adult mortality and is more likely to be associated with the size of the centenarian population. However, it can be more affected by data errors at later ages, and it is not available from 1900 to 1930.

In addition to the estimates based on CD Model Life Tables, we calculated survival functions by fitting three mathematical functions to Brazilian mortality rates: Gompertz, Weibull, and Kannisto. There is a general agreement that mortality increases exponentially from mid-adult to ages 80-90, as described by the Gompertz law. However, there is not a consensus regarding the mortality trajectory at the most advanced ages (Barbi et al. 2018; Gampe 2010; Gavrilova and Gavrilov 2015; Gavrilov and Gavrilova 2011; Robine and Vaupel 2001; Vaupel et al. 1998). Some studies suggest that the exponential growth of mortality with age is followed by a period of deceleration, with slower rates of mortality increase at the oldest ages (Barbi et al. 2017; Gampe 2010; Robine and Vaupel 2001; Vaupel et al. 1998). Conversely, another group of researchers has contended that mortality deceleration in later life is a consequence of poorer data quality at older ages, and that in reality mortality continues to grow exponentially at the highest ages (Gavrilov and Gavrilova 2011; Gavrilova and Gavrilov

2015). Due to this disagreement, we decided to offer estimates for the number of centenarians in Brazil from the application of the three mortality models.

To estimate the parameters of the mathematical models we used official life tables estimated by the Brazilian Institute of Geography and Statistics (IBGE 2013). We fitted the mathematical models to the death rates for the ages 70 to 90+ and then extrapolated the results to the advanced ages. To avoid adding bias from mortality data errors, which were more prevalent in the earlier years, we limited the mathematical estimates to the year 2000.

Sensitivity Analysis

As we stressed before, to test for the robustness of our results and offer a range (upper and lower limits) of estimates, we ran sensitivity analyses by varying the inputs of the Eq.1. First, we changed age *y* by starting the calculations from ages 60-69, instead of just 50-59. As we varied age *y*, we had to obtain the corresponding population estimates from the census data and the life table survivorship ratios for each new age range. Second, in addition to applying the annualized mean growth rate for ages 90 and above, we tested for two different open-ended age intervals: 80 and above, and 100 and above.

After varying the lower limit age *y*, the configuration of age-specific growth rates in the open-ended age interval and the mortality schedules, we produced 184 estimates of the centenarian population by sex between 1900 and 2000. We present and discuss absolute sizes, sex ratios, and prevalence rates for the Brazilian centenarian population, as well as the life table survivorship ratios from ages 50-59 to 100-109 under the scenarios previously discussed. We compare some of them to the measures calculated directly from census data.

Cohort Estimates

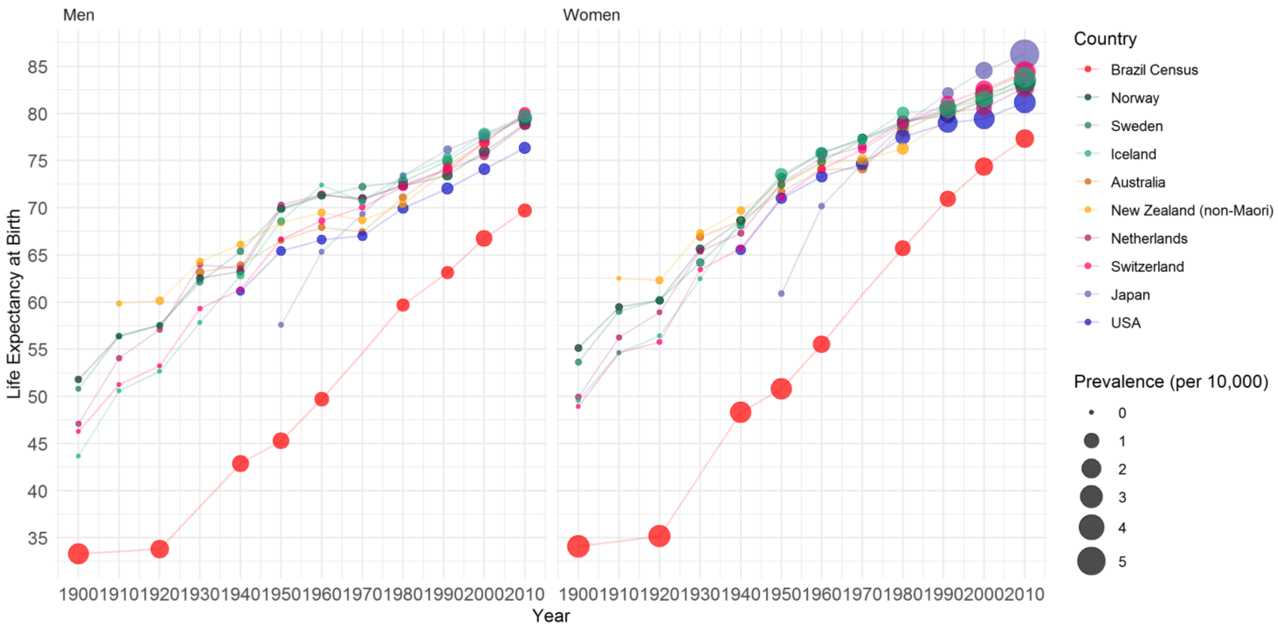
Variable-*r* estimates are useful to relate the number of centenarians to the number of younger adults in the same year, allowing us to estimate the population of Brazilian centenarians over a more extended period. However, there is no guarantee that the age-specific growth rates are not biased because of changes in census coverage and content errors over time. Therefore, we calculated another set of results based on cohort estimates for the second half of the twentieth century. We reconstructed the possible mortality trajectories for cohorts aged 50 to 59 and 60 to 69 years old in 1900 to 1950 and calculated the number of surviving individuals at ages 100-119 in every decade from 1950 to 2000. Because there are no census data in 1910 and 1930, we interpolated the population at ages 50 to 69 in those years based on the observations from 1900, 1920 and 1940 censuses. We limited our cohort analysis to the life tables calculated from CD Model Life Tables (West Model) since the objective is only to test for potential discrepancies between period and cohort estimations. In order to approximate the cohort mortality functions, we apply the geometric mean of life table survivorship ratios at the beginning and end of each intercensal period and take the corresponding age-period estimates.

3.3. Results

Before examining the indirect estimates of the centenarian population, it is essential to introduce the problem of over-enumeration of older people in Brazil. Fig.1 compares the prevalence rates (per 10,000) of centenarians recorded in the Brazilian censuses with prevalence rates for a set of high-longevity countries (HLC). The graph also displays the association between life expectancy at birth and the prevalence rates for each one of the countries. We drew population and mortality data for the HLC from the Human Mortality Database (HMD, 2019).

As expected, life expectancy was lower in Brazil than in the HLC, for both men and women. However, particularly in the past, the prevalence rates of centenarians in Brazil were considerably higher than in the HLC. For instance, there was virtually no centenarian in Sweden in 1900 and 1910, while in Brazil, the prevalence rate reached levels higher than two per 10,000 in 1900, according to the official figures. Over the twentieth century, as mortality declined, the prevalence of centenarians consistently increased in the HLC. In Sweden, for example, the prevalence of female centenarians augmented by about one hundred-fold over the 110 years, and in Japan, it increased by about three hundred-fold between 1950 and 2010. The prevalence rates for women in these two countries are now above two (Sweden) and five (Japan) per 10,000. At the same time, according to the Brazilian official statistics, the prevalence of centenarians took the opposite direction in Brazil. In the case of women, it decreased from 2.83 in 1900 to 1.75 per 10,000 in 2010, despite an increase of about 40 years in life expectancy at birth. For men, the decline in the prevalence rate was more substantial, ranging from 2.32 to 0.78 per 10,000.

Fig.1. Life expectancy at birth and prevalence of centenarians: Brazil and selected countries, men and women, 1900-2010



Source: HMD (2019) and Brazilian Census Data

The historical analysis of Brazilian census data raises serious questions about the quality of enumeration at older ages and suggests the need for new statistics of the centenarian population. As aforementioned, we produced 184 different indirect estimates, which are available upon request, based on the combination of several mortality models and parameters for Eq.1. Table 1 below summarizes the number of centenarians by sex, for each year, by showing the average, minimum, and maximum population calculated from the main set of indirect estimates. We cannot show the official figures for 1970 because the census data available in that year preclude us from separating the ninety-year-olds from centenarians.

Table 1. Number of centenarians in Brazil: census and indirect estimates, 1900-2000.

Year	Men				Women				Sex Ratio	
	Census	Estimated			Census	Estimated			Census	Estimated
		Average	Minimum	Maximum		Average	Minimum	Maximum		
1900	2,047	0	0	1	2,391	1	0	1	0.86	0.62
1920	2,625	1	0	2	4,102	2	0	3	0.64	0.46
1940	2,817	5	0	11	4,999	21	3	51	0.56	0.22
1950	3,290	7	2	13	6,399	22	5	46	0.51	0.33
1960	3,483	12	3	21	6,393	33	9	66	0.54	0.38
1970	-	22	5	48	-	65	17	146	-	0.34
1980	3,545	35	7	99	6,705	103	27	256	0.53	0.34
1990	4,253	48	15	115	8,773	168	54	353	0.48	0.29
2000	10,423	180	38	393	14,153	739	198	1,334	0.74	0.24

Note: The Brazilian Bureau of Census has not published the number of centenarians for 1970, only the population 90 years and older.

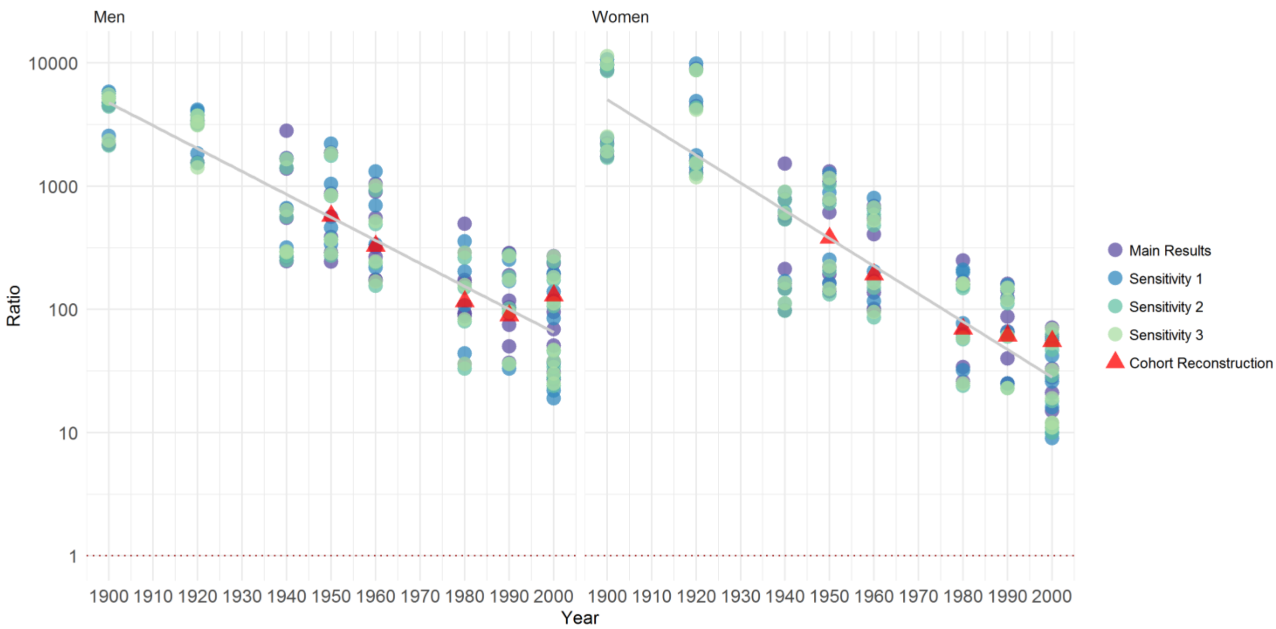
Source: Census Data. Calculations by the authors

The indirect estimates indicate that there was probably no centenarian in the first two decades of the twentieth century in Brazil. The number of centenarians grew little until 1940 when the mortality transition started to affect both population growth rates and the chances to reach ages 100 and older. It would have been only after 1960 that the number of centenarians had surpassed 100 individuals, reaching some value around 1,000 (between 200 and 2,000) centenarians in 2000, of which approximately 80% were women. Between 1900 and 2000, the sex ratio based on our estimates considerably declined from 0.62 to 0.24, because of an increasing gender gap in mortality over the years.

The estimates also confirm an extensive over-enumeration in the census records (Table 1). According to the census data, 4,438 centenarians were living in Brazil in 1900, from which 46% were men. This census estimate is so large that it was not even reached a century later, according to our most optimistic scenario. Also, the sex ratio of 0.86 in the 1900 census is surprising given the female mortality advantage, particularly at adult ages. This result shows that in the past, the over-enumeration of older groups was more frequent among men than women.

The implausibility of the census data becomes even more evident in Fig.2. We calculated ratios of the centenarian population in the census data to all indirect measures based on the variable-r method and the cohort estimates. The declining ratios for both men and women confirm the implausible numbers in the past and the improvement of census data over the years. In the first decades of the twentieth century, prevalence rates based on indirect estimates were a few thousand times lower than the ones calculated from the censuses. In 2000, the mean ratio (census data/estimates) was significantly lower than in the past but still higher than 10.

Fig.2. Ratio of the number of centenarians recorded in the census and indirectly estimated according to different parameters: Brazil, men and women, 1900−2000



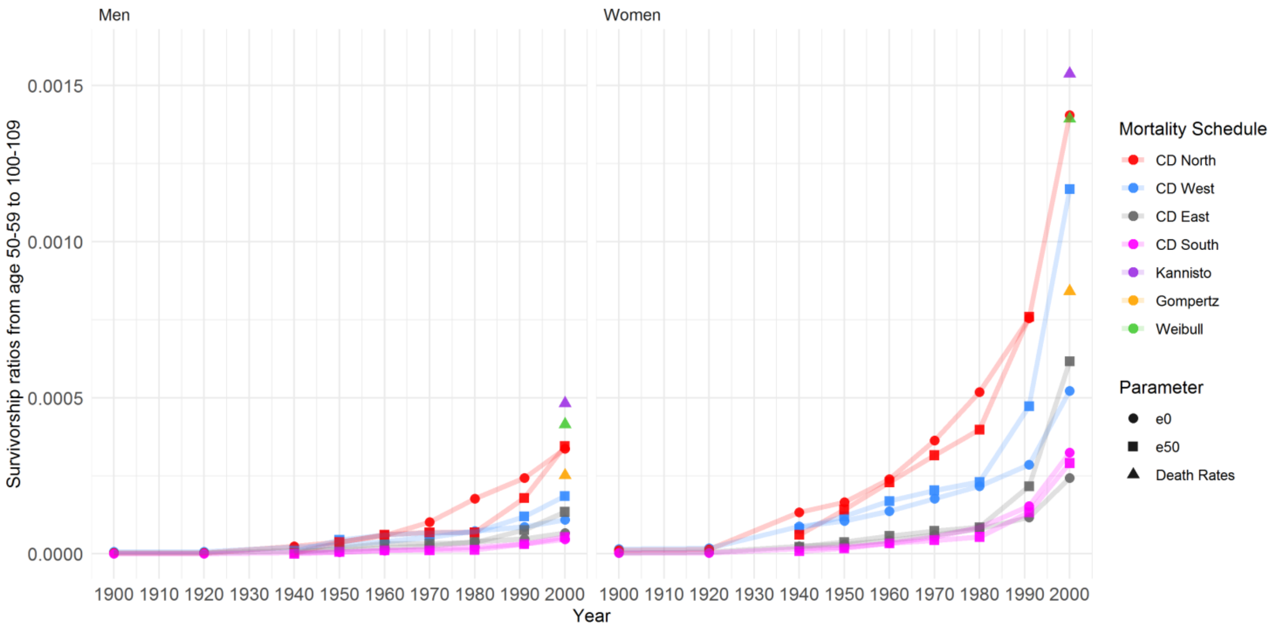
Notes: The main results and sensitivity analysis are based on the variable-r method. The lower limit age group is equal to 50-59 for the main results, and both sensitivity 1 and 2. It is equal to 60-69 for sensitivity 3. The open age interval for the calculation of the growth rates is equal to 90+ for both the main results and sensitivity 3, and equal to 80+ for sensitivity 1 and 100+ for sensitivity 2. The dotted line corresponds to an equal number of centenarians recorded in census and indirectly estimated.

Source: Census Data. Calculations by the authors

Indirect estimates of the centenarian population also involve uncertainty. Fig.2 reveals the sensitivity of the centenarian population indirectly estimated due to changes in the parameters of Eq.1. Changing the age y (lower age limit) from 50-59 to 60-69 or varying the open-ended age group (from 90+ to 80+ and 100+) to calculate the age-specific growth rates have only minor impacts on the estimated number of centenarians. Depending on the parameter changed, the average relative difference calculated for all the years and mortality models together varies from less than four to about 19% among men, and from four to about to about 15% among women. Consistently, cohort reconstructions based on the West family of the CD Life Tables are also comparable to the variable-r estimates. Among men, from 1950 to 2000, the average number of male centenarians would be only 11% higher using period instead of cohort measures. Among women, the average difference is only about 6%, for the years 1950 to 1990. However, for the year 2000, there is a more considerable discrepancy. We estimated 425 and 258 women, respectively, in the period and cohort analyses. This isolated result may reflect a combination of factors, including the inadequacy of the life table model, the increasing gender differences in mortality in more recent years, and changes in census coverage at older ages between 2000 and 2010 that impacted the age-specific growth rates.

By analyzing Table 1 and Fig.2 together, we see that the most substantial portion of the variation in our results comes from the choice of the adult mortality model. For instance, in 1960, the number of centenarians can vary from 12 to 87 only due to changes in the mortality schedules. Therefore, the idea of providing a set of population estimates by varying the mortality models is not only to uncover our ignorance about the actual mortality function at older ages in Brazil but also to measure the extent to which census data are out of range. In Fig.3, we show how life table survivorship ratios from ages 50-59 to 100-109 have changed since 1900, according to each mortality age schedule. The ratios derived from the North, Kannisto, and Weibull model are the highest, since (i) the North Model reflects conditions that are typical of countries that experience lower mortality at older ages, (ii) the Kannisto Model follows the logistic form at advanced ages, and (iii) the Weibull Model is an intermediate pattern between the exponential and logistic mortality increases at older ages. The lowest survival estimates are those calculated from the South and East Model, which is not surprising because they reflect conditions of countries with higher mortality at advanced ages. The difference in the ratios estimated between the highest and the lowest models can get as high as ten times and seven times, respectively, for men and women in 2000.

Fig.3. Life table survivorship ratios calculated from selected mortality schedules: Brazil, men and women, 1900−2000



Source: Calculations by the authors

Regardless of the mortality model adopted, survival gains have been particularly high among women. For example, the chances of women surviving to ages 100-109 from ages 50-59 improved 34-fold, between 1900 and 2000, according to the West Model. However, these gains were not enough to explain the large number of centenarians in the Brazilian censuses. The actual probabilities of surviving would need to be ten to one thousand times higher than the survival curves adopted in the current work, depending on the year and sex, for census records to be real.

3.4. Discussion

There is a general belief in Brazil that the centenarian population has been increasing very rapidly over the last decades. According to the census data, the centenarian population grew two and a half times between 1960 and 2000 and currently comprises more than 24,000 individuals. Whereas the Brazilian population is indeed getting older faster than in many wealthier regions, the estimated centenarian population is probably significantly smaller than the number of individuals counted by the census, having increased from fewer than 100 individuals in 1960 to between 200 and 2,000 people in the year 2000. The slower growth of the centenarian population in the census than predicted by our estimates suggests there have been improvements in data collection over time. Also, the much higher census sex ratios reinforce the hypothesis of over-enumeration of centenarians, particularly men.

Our indirect estimates based on variable-r relations are not immune to errors. Age-specific growth rates in Brazil may vary over time due to changes in census data quality, affecting the estimates of the population size at older ages. However, sensitivity analyses, including cohort estimates, indicated that any errors from the methodology are probably not substantial. What affects population estimates at old ages is the choice of the mortality model. To minimize the lack of high-quality mortality data at advanced ages in Brazil, we compared estimates from different mortality schedules and concluded that unless adult age patterns of mortality are very atypical in Brazil, the number of centenarians in the census records are incompatible with the prevailing adult mortality levels in the country.

Mortara (1949) was the first demographer to worry about the poor quality of the population enumerated at older ages in the Brazilian censuses, anticipating what would be a crucial subject decades later. Ours is the first study to provide a systematic comparison between indirect estimates and census records of centenarians over a century in Brazil, although other recent studies have also cast light on issues related to the enumeration of centenarians. As aforementioned, Gomes and Turra (2009) found that the number of centenarians in census records was triple those in death files in the year 1991. Here, we showed that the number of centenarians is not only substantially smaller than census records but also lower than the number of deaths registered at ages 100 and older. Whereas it is not surprising that age misreporting affects the quality of both census data and death files, the bias is proportionally more extensive in the census records resulting in artificially lower mortality rates at older ages, as recognized by (Preston, Elo, and Stewart 1999) and examined by Turra (2012) for Brazil.

The evolution of the population of Brazilian centenarians across the decades, as measured by the census records, is odd, because of significant data improvements that have occurred in each new data collection. This inaccurate time trend seems to reflect in the estimates prepared by the U.N. (2019). According to them, the number of centenarians decreased from almost nine thousand in 1950 to fewer than 100 individuals in 1970, increasing after that. Therefore, we believe the U.N. figures for the 1950s and 1960s are wrong, although they are significantly more consistent with (somewhat higher than) our estimates for the period 1970-2000. Also, according to the U.N. (2019) projections, the population of Brazilian centenarians will only reach the number recorded in the 2000 census (about 25,000 individuals) in the year 2025, surpassing 100,000 and 1,000,000 cases, respectively, after 2035 and 2085. The estimation of future trends still needs more scrutiny, since the correct age patterns of mortality at adult ages in Brazil remain a puzzle. Fortunately, we expect the quality of age reporting at older ages to improve over the next decades, with the educational transition and the consolidation of the vital registration systems. The accurate measurement of population and deaths at adults ages in Brazil will become a much smaller problem in a few decades.

4. Assessing the quality of self-reported education among adults in Brazil, with intercensal survivorship ratios

4.1. Introduction

Education is one of the keys to development and economic growth and is connected to demographic variables in many different ways. In virtually all societies, individuals with higher levels of education enjoy better health and longer lives (Preston and Taubman 1994; Elo and Preston 1996; Mackenbach et al. 1999; Koch et al. 2007; Zhu and Xie 2007; Rosero-Bixby and Dow 2009; Turra et al. 2016; Lutz and Kebede 2018; Smith-Greenaway and Yeatman 2019). Highly educated couples have fewer children, mainly because they marry later, use contraception more effectively, and have more autonomy in reproductive decision-making (Singh and Casterline 1985; Potter et al. 2010; Bongaarts et al. 2017; Rios-Neto et al. 2018). Moreover, migration flows across regions with different developmental levels are usually associated with selection of migrants by education group (de Haas 2010; Lewis 1986). It is thus not surprising that as educational enrolments expand worldwide, there is a growing emphasis on population forecasts by educational level (KC et al. 2010; Lutz and KC 2011; Lutz et al. 2014).

As an essential variable in demographic analysis, accurate measurement of education is indispensable. Studies of the quality of reported education are not new in the literature (Folger and Nam 1964; Gustavus and Nam 1968; Black et al. 2003; Sorlie and Johnson 1996; Kane et al. 1999; Johnson-Greene et al. 1997; Battistin et al. 2014; Lerch et al. 2017). Some of these have documented the misreporting of educational levels in data systems. For example, in both the 1950 and 1960 U.S. census, there was a significant amount of net overreporting, particularly at older ages for high school and college levels (Folger and Nam 1964). Examination of the 1990 U.S. census detected misreporting of some doctoral and professional degree categories (Black et al. 2003).

One of the consequences of education misreporting in census data is the miscalculation of the denominator of demographic rates, which can bias the measurement of educational gradients. In mortality analysis, for example, where death rates usually derive from two independent data sources - vital statistics (death counts) and census data (exposures-at-risk) – disagreement between reported education in the numerator and denominator may affect conclusions about how mortality differentials by education vary across age, gender, race, and geographical groups. Several authors have documented disagreements between death counts and population census records by education in Europe and the U.S.. The level of disagreement tends to increase with age in Lithuania (Shkolnikov et al. 2007), and for some racial/ethnic groups in the U.S. such as blacks and Hispanics (Rostron et al. 2010). However, even among population groups that are less likely to report education levels differently across data sources, changes in the census questions can lead to inconsistent information over time (Lerch et al. 2017).

In Brazil, research has neglected the possible effects of education misreporting in demographic, social, and economic studies. Although several authors have looked at coverage errors, both in vital and census data (Paes and Albuquerque 1999; Paes 2007; Lima and Queiroz 2014; Cavenaghi and Alves 2016), demographers have paid less attention to content errors, and analysis has mostly been limited to age misreporting (e.g. Paes and Albuquerque 1999; Agostinho 2009; Gomes and Turra 2009; Turra 2012; di Lego et al. 2017; Nepomuceno and Turra 2019). Nevertheless, the high degree of inaccuracy in age reporting suggests that problems of misreporting may affect other variables, including education.

This article examines the quality of education information in Brazilian census data. To do this, we calculate mortality levels by education as implied by intercensal survivorship ratios, to investigate the quality of self-reported education data among adults in Brazil between the 1991 and 2000 censuses. Here, we hypothesize that measurement errors are more substantial during times of accelerated expansion of schooling, such as the period that characterized the education system in Brazil after 1980 (IBGE 2004; Rios-Neto et al. 2010). The rapid education transition may have influenced the reliability of education data in at least three different ways. First, if better-educated individuals tend to report their characteristics more accurately, mainly when information is retrospective, errors will be more substantial in earlier data collections and among older cohorts. Second, the expansion of education may have changed people's perception of their relative social position. Older age groups, who lived their schooling years before the development of the education system, may feel inclined to overstate their educational levels to level-off any cohort differences. Finally, the expansion of schooling has been followed by changes in the educational system and census questions, affecting the comparability of responses over time. Given the global effort to improve access to schooling (United Nations 2015, 2016), our results aim to draw attention to the accuracy of reported education not only in Brazil but also in those countries where educational expansion is underway, and deficient data quality is a potential issue.

4.2. Data and Methods

We drew data for men and women 40–89 years old, from the 1991 and 2000 Brazilian Census (IPUMS-I 2019). Although the most recent census data available are from 2010, the exclusion from the questionnaire of information on the highest grade completed precluded us from using the data for that year. To have an exact 10-year interval between periods, we estimated the population for the year 1990 based on the set of age-specific growth rates between 1991 and 2000 and the population in 1991. We selected only individuals who were not attending school at the census time, representing about 99.5% of the population aged 40 to 89 in 1990 and 97.3% in 2000.

We measure educational attainment as the highest grade completed within the most advanced level attended in the education system. We used the IPUMS' harmonized variable named "YRSCHOOL" which accounts for the number of years of schooling.

To evaluate the quality of self-reported education, we assessed the implicit mortality by education between the two censuses through intercensal survivorship ratios (ISR) by age, sex, and educational attainment:

$$ISR_x^{k,i}(j) = \frac{N_{x+j}^{k,i}(t+j)}{N_x^{k,i}(t)}$$

where $N_x^{k,i}(t)$ is the cohort at age *x*, sex *k*, at educational level *i* in the year *t*, and $N_{x+j}^{k,i}(t+j)$ is the same cohort *j* years older at the educational level *i* and sex *k* in the year *t* + *j*. In our first analysis, we used single year of schooling, from 0 to 12+. Next, we calculated the number of years of schooling categorized into four intervals: 0–3, 4–8, 9–11, and 12+, that correspond respectively to the first and second stages of the primary, secondary and tertiary education.

After the categorization of the years of schooling, values of the ISR at each age were translated into life expectancy at age 40 in the Coale-Demeny West model life-table (Coale et al. 1983). We used the United Nations version of the Coale-Demeny Model

Life Tables, which extends the original mortality levels to include life expectancy at birth up to 100 years (United Nations 2017). This allowed us to assess the consistency of the adult level of mortality across age and educational groups.

The ISR capture changes in the cohort size during the intercensal period. In the absence of international migration and changes in census data quality, one should expect survivorship ratios to be less than one due to the impact of mortality. Here we assume that the population was closed to international migration at ages above 40, between 1990 and 2000. This assumption is reasonable because the international net migration rate was very low during the 1990s in Brazil, reaching less than half of one percent for the population aged ten years and older (Carvalho and Campos 2006; Campos 2011), and an even lower values at older ages (Garcia 2013). To mitigate the effect of changes in census coverage on the ISR, we also adjusted the census enumeration according to omission rates (1.04 and 1.02, respectively for males and females for 1990, and 1.03 for males and 1.01 for females in 2000) published by the Brazilian Institute of Geography and Statistics (IBGE 2003).

We expect survivorship ratios to increase with educational attainment. Earlier analyses that used surveys or data from the Mortality Information System of the Ministry of Health suggested a strong educational gradient in adult mortality in Brazil (Renter´ıa and Turra 2009; Turra et al. 2016, 2018). Despite the existence of public programs to improve or supplement adult education, we expect only negligible gains in education over time, at ages above 40 in Brazil. Between the 1991 and 2000 censuses, the proportion of individuals who reported not attending school varied from 99.2% to 96.5% for the age group 40–49, and from 99.6% to 98.2% for the age group 50–59. For older age groups, these proportions were even higher and more similar in the two censuses: between 99.7% and 99% for the age group 60-69, 99.8% to 99.3% for the age group 70–79, and 99.8% and 99.4% for individuals aged 80 and older.

4.3. Results

Table 1 presents the education distribution by age groups and sex in Brazil. In both years, there is a substantial proportion of adults in the least educated groups, and a smaller percentage in the most educated categories. Between 1990 and 2000, the most considerable change in the education distribution is among the least educated (0–3 years of schooling), mainly for younger age groups because of the advance of the education transition in the country. For older age groups, proportions change only slightly over time.

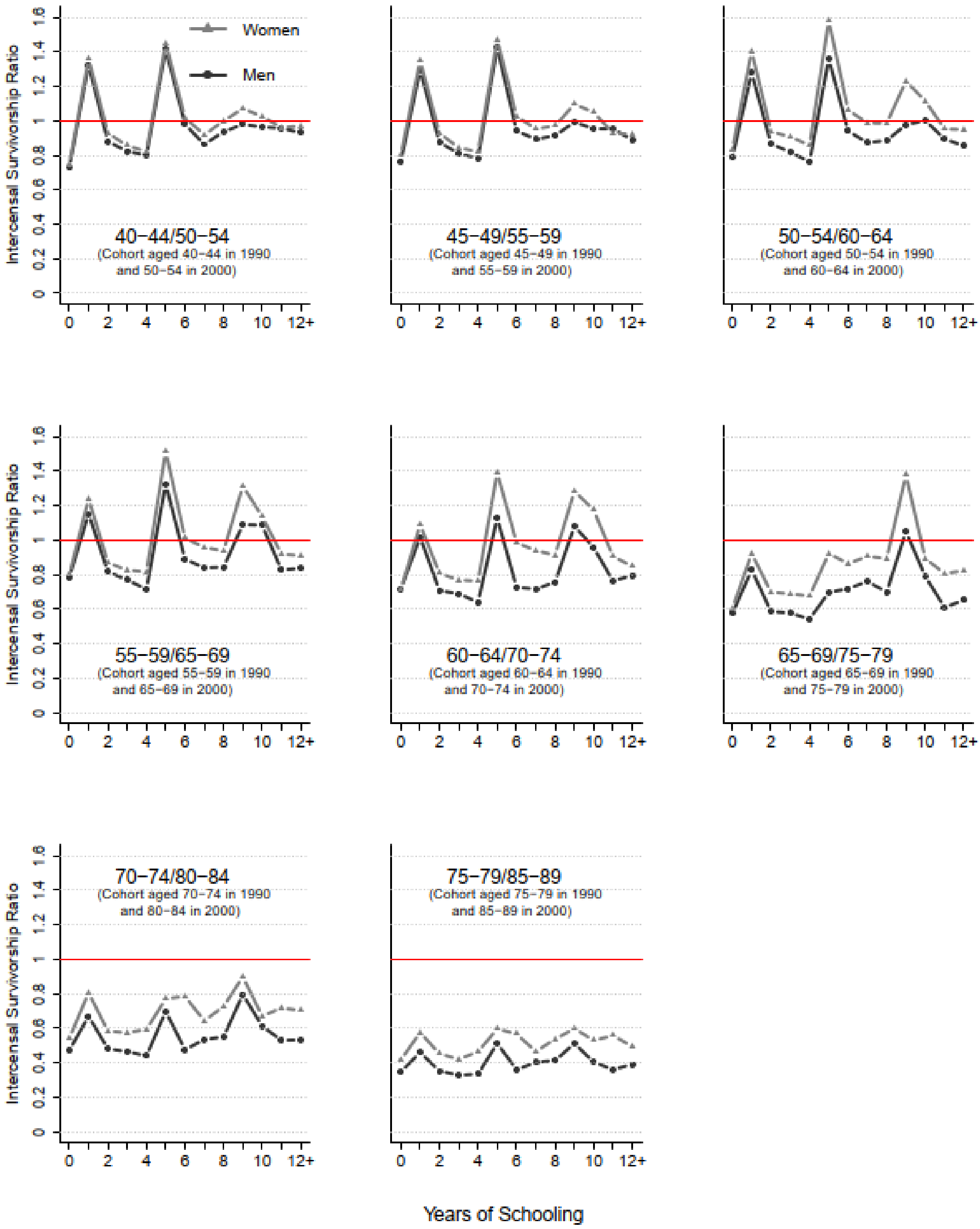
Table1- Education distribution (%) by age: Brazil, men and women 1990 and 2000.

Age Group	0-3		4-8		9-11		12+	
	1990	2000	1990	2000	1990	2000	1990	2000
Men								
40-49	42.95%	29.35%	38.39%	42.83%	9.71%	16.95%	8.95%	10.87%
50-59	52.43%	41.69%	34.78%	38.13%	6.80%	10.73%	5.98%	9.45%
60-69	61.40%	53.25%	29.34%	33.33%	4.76%	7.29%	4.51%	6.13%
70-79	69.86%	61.95%	23.50%	27.90%	3.09%	5.18%	3.55%	4.97%
80-89	73.98%	69.87%	20.45%	22.85%	2.55%	3.40%	3.03%	3.88%
Women								
40-49	46.05%	29.12%	36.40%	41.59%	10.31%	18.11%	7.24%	11.18%
50-59	57.30%	44.67%	32.49%	36.55%	6.55%	11.07%	3.66%	7.71%
60-69	65.01%	56.56%	28.14%	32.55%	4.79%	7.06%	2.06%	3.84%
70-79	71.10%	62.41%	23.59%	29.48%	4.13%	5.76%	1.18%	2.34%
80-89	74.09%	68.32%	21.71%	25.18%	3.40%	5.11%	0.80%	1.40%

Source: Census data (IPUMS-I 2019)

Figure 1 shows intercensal survivorship ratios by age, sex, and single year of schooling. The estimates show an irregular and unexpected pattern of the ISR by education. First, except for the oldest cohorts (70–74 and 75–79 years old in 1990), the ISR can get higher than one, suggesting that the size of cohorts increased between the censuses for some years of completed schooling, despite the impact of mortality. Second, the ISR do not increase monotonically with education. Although survivorship tends to increase at higher years of schooling, it fluctuates over the distribution, varying sharply, particularly at one, five and nine years of schooling. For example, the number of men and women aged 40–44 with one year of schooling increased by more than 30% between 1990 and 2000, and more than 40% for those with five years of schooling. For nine years of schooling, the number of women and men aged 55–59 reporting this level increased by 30% and 10% respectively between 1990 and 2000.

Figure 1: Intercensal Survivorship Ratio by age and years of schooling: Brazil, men and women, 1990-2000.



Source: Census data (IPUMS-I 2019)

One way to re-examine education reporting is by categorizing years of schooling into intervals (0–3, 4–8, 9–11, and 12+). Table 2 shows the results, including the ISR and the life expectancy at age 40 (e40) from the West Model of the Coale-Demeny Life Tables implied by the ISR.

To further improve our estimates, we also re-categorized the 5-year into 10-year age intervals, hoping to minimize the possible effects of age misreporting in the census, particularly at older ages (Agostinho 2009; Nepomuceno and Turra 2019).

Table 2 shows the educational gradient in mortality, i.e., the positive relationship between education and survival. Life expectancy at age 40 is 0.44 to 7.84 years lower for individuals with 0–3 years of schooling than for those with 12 or more years of schooling, depending on the cohort and sex.

Some implausible results remain after the categorization of years of schooling. Table 2 reveals that the survival advantage of the most educated does not decrease monotonically with age. For instance, the difference in e40 between the least and the most educated drops from 7.8 years for women aged 40–49 to 5.2 to those aged 50–59 in 1990. But it then increases to 6.1 years at the age group 60–69, reducing to 4.0 at 70–79 years in 1990. By comparing the individuals with 9–11 and those with 12 years or more of schooling, we identify other implausible results that cast doubt on the quality of self-reported education in the Brazilian census. For most male and female cohorts, life expectancy reduces for those with 12+ years of schooling compared with 9–11 years of schooling, contradicting the expected gradient. Table 2 also shows that mortality levels implied by the ISR vary significantly across cohorts/age. For all educational categories and both sexes, the e40 implied by the ISR increases for the older cohorts, and the difference can get as high as twelve years for the lowest educated group.

Table2- Intercensal Survivorshop Ratio by educational groups and decennial age groups: Brazil, men and women

Age in 1990	Age in 2000	0-3		4-8		9-11		12+	
		Survival Ratio	Implied e_{40}	Survival Ratio	Implied e_{40}	Survival Ratio	Implied e_{40}	Survival Ratio	Implied e_{40}
Men									
40-49	50-59	0.8410	26.66	0.8603	27.94	0.9570	36.21	0.9145	31.64
50-59	60-69	0.8398	33.52	0.7925	30.67	0.8867	36.79	0.8469	33.96
60-69	70-79	0.6747	34.94	0.6358	33.17	0.7287	37.26	0.7381	37.56
70-79	80-89	0.4350	37.46	0.4229	37.70	0.4785	38.46	0.4762	38.35
Women									
40-49	50-59	0.8671	28.51	0.8974	30.76	0.9600	37.64	0.9523	36.34
50-59	60-69	0.8798	35.51	0.8931	36.49	0.9607	44.30	0.9344	40.74
60-69	70-79	0.7090	35.87	0.7739	38.74	0.8881	45.15	0.8388	41.95
70-79	80-89	0.5082	40.20	0.5646	42.04	0.6545	44.94	0.6241	44.20

Note : life expectancy at age 40 from the Model West implied by the ISR.
Source: Census data (IPUMS-I 2019)

4.4. Discussion and Conclusion

Our study shows evidence of education misreporting in Brazil. The ISR by single year of schooling only barely reflects the known educational gradient in mortality described in the literature. After the categorization of the years of schooling, the positive relationship between education and survival shows up, although some implausible results remained.

An ISR higher than one, mainly at one, five and nine years of education, suggests an unlikely increase in the cohort size between censuses. In the Brazilian education system, these grades corresponded to the beginning of stages 1 and 2 of primary (1–4, 5–8 years of schooling) and the beginning of secondary (9–11 years of schooling) in 1990. Therefore, the implausible high ISR at these points of the distribution suggests two possible (complementary) explanations. First, some adults may have genuinely moved between levels of educational attainment over the intercensal period, more than compensating the mortality effect. Second (and more likely), a fraction of the cohorts may have misclassified themselves as literate or having some stage 2 of primary or secondary education in the second census, because of memory error and other self-reporting inconsistencies caused by changes in the social context, the census questionnaire, and education reforms.

The cohorts we analyzed experienced at least two education reforms after their schooling years that greatly changed the structure and the terminology of the education system, one in 1971, and another in 1996. The first reform, for instance, merged two formerly levels that corresponded, respectively, to 1–4 grades and 5–8 grades into a new level with 1–8 grades (Rigotti and Cerqueira 2004). Individuals who lived their schooling years before these reforms may have problems to classify themselves according to the new education system, resulting in misclassification of their educational level. Further, census questions to measure educational attainment changed over time, which may be another source of inconsistency. In 1991 census, the questions to calculate the highest grade completed within the most advanced level attended were "What was the last grade passed?" and "What was the last level passed"; while in 2000 the questions were "What was the highest course attended, in which you completed at least one grade?" and "What was the last grade passed?". These mentioned changes point out some challenges in the comparison of the years of schooling over time, by age and cohorts, as revealed by our findings.

Education gains may also contribute to the implausible higher ISR. Since the number of individuals aged 40–49 with one year of schooling increased by more than 30% during the intercensal period, and part of this may be due to education gains, we checked the proportion of illiterates who were attending school in the 1991 census, and could potentially achieve one year of schooling between 1991 and 2000. This proportion was very low, reaching less than 0.05% in 1991, and less than 0.06% in the 2000 census. This small proportion is not enough to explain the striking increase in the population aged 40—49 with one year of schooling between the censuses.

To minimize the effects of age and education misreporting, we presented results by 10-year age groups and four intervals of years of schooling. After the re-categorization, the educational gradient in mortality became more explicit, and the estimated ISR were all lower than one. The categorization of the years of schooling is an alternative to reduce the effect of education misreporting. However, since there are substantial socio and economic differences among individuals within groups, by categorizing, we cannot measure important educational differences in mortality due to the inconsistencies revealed by the ISR by single years of schooling.

Our results also show some inconsistencies after the re-categorization, including lower ISR for individuals with 12+ than 9–11 years of schooling. One important finding is the lack of consistency in mortality levels by age. According to the e40 implied by the ISR, older cohorts experienced lower levels of mortality at older ages. The difference is particularly strong for the least educated. The lower levels of mortality at older ages could be explained by the unusual adult age pattern of mortality in Brazil, which has been shown to increase with age at a slower rate than in countries with high-quality population and mortality data (Turra 2012).

Since the estimates provided here show evidence of education misreporting in the census, our findings increase concerns about the true educational distribution of the adult population. Further, since the data on education attainment seem to be differentially misreported by age and educational level, the validity of age-specific demographic rates by educational level should be interpreted cautiously in Brazil. Furthermore, if education misreporting is different across censuses, trend analysis using these data will reflect erroneous patterns. Global studies that use census education data from developing countries to project population should be aware of this weakness. Lastly, our findings draw attention to the importance of investigating the potential bias in demographic rates by educational levels in Brazil and in other developing countries where educational expansion is underway.

This study is just the first step in revealing potential errors in census education data, and we still do not know nearly enough. Efforts need to be made to measure the magnitude and the direction (overreporting and underreporting) of misreporting of census education data, for the whole population and for subgroups. Different sources of education data could be used to estimate the amount of education misreporting and provide adjusted figures. Linkage with administrative records may help, although this type of data is rarer at older ages. At the same time, attempts to reduce misreporting of education should rely on improvements in census data collection, such as in phrasing of relevant and comparable questions over time and the reduction of omission rates.

5. Final Remarks

This PhD dissertation dealt with data quality issues that are related, directly or indirectly, with survival analysis. The first paper presented an age-cohort decomposition of mortality gap that it is not limited to populations with complete cohort mortality data. We applied an alternative approach, *The Truncated Cross-sectional Average Length of Life* (*TCAL*), which enables us to identify specific cohorts that account for most of the disparity in mortality between the populations. The second paper provided indirect estimates of centenarian population, for a period longer than a century in Brazil. The comparison between our estimates and census records of centenarians suggested that centenarians are over-enumerated in census. Thus, we concluded that age misreporting greatly affect the quality of the oldest-old census population. The final article also focused on analysis on content errors, but in the education data, giving clear evidences of education data inaccuracy in adult population. In conclusion, this dissertation, provide alternative ways to minimize the lack of complete cohort mortality data or the existence of inaccurate data.

References

Agostinho, C. S. (2009). Estudo sobre a mortalidade adulta, para Brasil entre 1980 e 2000 e Unidades da Federacao em 2000: uma aplicacao dos metodos de distribuicao de mosres. PhD thesis, Universidade Federal de Minas Gerais, Faculdade de Ciencias Economicas Cedeplar.

ANDREEV E., NOLTE E., SHKOLNIKOV V., VARAVIKOVA E., MCKEE M., 2003, The evolving pattern of avoidable mortality in Russia, International Journal of Epidemiology 32 (3), 437–446.

Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W., & Wachter, K. W. (2018). Demography of longevity pioneers. Science, 360(6396), 1459–1461.

Battistin, E., Nadai, M. D., and Sianesi, B. (2014). Misreported schooling, multiple measures and returns toeducational qualifications. Journal of Econometrics, 181(2):136–150.

BERGERON-BOUCHER M., EBELING M., CANUDAS-ROMO V., 2015, Decomposition changes in life expectancy: Compression versus shifting mortality, Demographic Research (33) 14, 391–424.

Black, D., Sanders, S., and Taylor, L. (2003). Measurement of higher education in the census and current population survey. Journal of the American Statistical Association, 98(463):545–554.

BOBAK M., 1996, Contribution of socioeconomic/psychosocial environment–Evidence from studies within Central and Eastern Europe, in Hertzman C., Shona K., Bobak M. (eds.), East–West life expectancy gap in Europe: Environmental and non-environmental determinants, Dordrecht, Springer Netherlands, 121–131.

BOBAK M., MARMOT M., 1996, East–West mortality divide and its potential explanations: Proposed research agenda, British Medical Journal 312 (7028), 421–5.

Bongaarts, J., Mensch, B. S., and Blanc, A. K. (2017). Trends in the age at reproductive transitions in the developing world: The role of education. Population Studies, 71(2):139–154.

Bourbeau, R., & Lebel, A. (2000). Mortality statistics for the oldest-old. Demographic Research, 2(2),1-36.

BRASIL, Ministerio da Agricultura, Industria e Commercio (1927). Annuario estatistico do Brazil 1908-1912 [Brazilian Statistical Yearbook 1908-1912]. Rio de Janeiro: Directoria Geral de Estatistica, (1-3), 1916-1927.

BRASIL, Ministerio da Agricultura, Industria e Commercio (1928). Recenseamento do Brazil realizado em 1 de setembro de 1920: população [Brazilian census conducted in 1st of September 1920]. Rio de Janeiro: Directoria Geral de Estatistica, 4(2), 1-925.

BROUARD, N. 1986, Structure et dynamique des populations. La pyramide des années à vivre, aspects nationaux et exemples régionaux. Espace Populations Societes, 4(2), 157-168.

Calzadilla, E. C., Fonseca, G.M, Borroto, M. & Martinez, C.P. (2013). Characterization of the nutritional status of the Cuban centenarian population. Revista Cubana de Investigaciones Biomédicas, 31(4), 459–466.

Campos, M. B. (2011). Reversao do saldo migratorio international negativo do Brasil? Evidencias preliminares com base nos dados do censo 2010. Revista Paranaense de Desenvolvimento Economico e Social, (121).

CANUDAS-ROMO V. , GUILLOT M., 2015, Truncated cross-sectional average length of life: A measure for comparing the mortality history of cohorts, Population Studies 69 (2), 147–159.

CANUDAS-ROMO V., 2010, Three measures of longevity: Time trends and record values, Demography 47 (2), 299–312.

CANUDAS-ROMO V., ENGELMAN M., 2012, The lagging behind of the US life expectancy: International and domestic comparison, Genus 68 (3), 1–22.

Carvalho, J. A. M. and Campos, M. B. (2006). A variacao do saldo migratorio internacional do Brazil. Estudos Avancados, 20(57):55–58.

Cavenaghi, S. M. and Alves, J. E. D. (2016). Qualidade das informacoes sobre fecundidade no censo demografico de 2010. Revista Brasileira de Estudos de Populacao, 33(1):189–206.

Coale, A. J., & Kisker, E. E. (1986). Mortality Crossovers: Reality or Bad Data. Population Studies, 40(3), 389–401.

Coale, A. J., Demeny, P., and Vaughan, B. (1983). Regional Model Life Tables and Stable Populations: Studies in Population. Academic Press.

de Haas, H. (2010). Migration and development: A theoretical perspective. International Migration Review, 44(1):227–264.

Dechter, A. R., & Preston, S. H. (1991). Age misreporting and its effects on adult mortality estimates in Latin America. Population Bulletin of the United Nations, (31–32), 1–16.

di Lego, V., Turra, C. M., & Cesar, C. (2017). Mortality selection among adults in Brazil: The survival advantage of Air Force officers. Demographic Research, 37(1), 1339–1350.

Drefahl, S., Lundström, H., Modig, K., & Ahlbom, A. (2012). The era of centenarians: mortality of the oldest old in Sweden. Journal of Internal Medicine, 272(1), 100–102.

Elo, I. T. and Preston, S. H. (1996). Educational differentials in mortality: United states, 1979–1985. Social Science & Medicine, 42(1):47–57.

Elo, Irma T., & Preston, S. H. (1994). "Estimating African-American Mortality from Inaccurate Data. Published by : Springer on Behalf of the Population Association of America, 31 (3), 427–58.

Folger, J. K. and Nam, C. B. (1964). Educational trends from census data. Demography, 1(1):247–57.

FORSTER M., 1996, The East–Westlife expectancy gap: The contribution of differences in the quality of medical care, in Hertzman C., Shona K., Bobak M. (eds.), East–Westlife expectancy gap in Europe: Environmental and non-environmental determinants, Dordrecht, Kluwer, 47–60.

Gampe, J. (2010). Human mortality beyond age 110. In: Meier H. et al. (Eds.). Supercentenarians. Demographic Research Monographs. Heidelberg, Springer.

Garcia, R. A. (2013). Estimativas dos emigrantes internacionais do brasil entre 1995 e 2000: uma aplicacao do metodo das razoes intercensitarias de sobrevivencia. Revisa Brasileira de Estudos de Populacao, 30(1):99–123.

Gavrilov, L. A., & Gavrilova, N. S. (2011). Mortality Measurement at Advanced Ages: A Study of the Social Security Administration Death Master File. North American Actuarial Journal, 15(3), 432–447.

GAVRILOVA N., EVDOKUSHKINA G., SEMYONOVA V., GAVRILOV L., 2001, Economic Crises, Stress and Mortality in Russia, paper presented at The Population Association of America 2001 Annual Meeting, Washington DC, March 28–31 2001.

Gavrilova, N. S., & Gavrilov, L. A. (2015). Biodemography of old-age mortality in humans and rodents. Journals of Gerontology - Series A Biological Sciences and Medical Sciences, 70(1), 1–9.

Gomes, M. M. F. and Turra, C. M. (2009). The number of centenarians in brazil: Indirect estimates based on death certificates. Demographic Research, 20:495–502.

Gomes, M. M. F., & Turra, C. M. (2009). The number of centenarians in Brazil: Indirect estimates based on death certificates. Demographic Research, 20(20), 495–502.

GRIGORIEV P., 2015, About mortality data for Belarus, The Human Mortality Database. Background and documentation.

GRIGORIEV P., ANDREEV E., 2015, The huge reduction in adult male mortality in Belarus and Russia: Is it attributable to anti-alcohol measures?, PLoS ONE 10 (9), 1–16.

GRIGORIEV P., MESLÉ F., SHKOLNIKOV V., ANDREEV E., AGNIESZKA F., PECHHOLDOV M., VALLIN J., 2014, The recent mortality decline in russia: Beginning of the cardiovascular revolution?, Population and Development Review 40 (1), 107–129.

GRIGORIEV P., SHKOLNIKOV V., ANDREEV E., JASILIONIS D., JDANOV DMITRI, MESLÉ F., VALLIN J., 2010, Mortality in Belarus, Lithuania, and Russia: Divergence in recent trends and possible explanations, European Journal of Population 26 (3), pp 245–274.

Gustavus, S. O. and Nam, C. B. (1968). Estimates of the “true” educational distribution of the adult population of the united states from 1910 to 1960. Demography, 5(1):410–421.

Herm, A., Cheung, S. L. K., & Poulain, M. (2012). Emergence of oldest old and centenarians: demographic analysis. Asian Journal of Gerontology & Geriatrics, 7(1), 19–26.

HO J. , PRESTON S., 2010, US mortality in an international context: Age variations, Population and Development Review 36 (4), 749–773.

Horiuchi, S. and Preston, S. H. (1988). "Age-Specific Growth Rates: The Legacy of Past Population Dynamics". Demography 25 (3): 429-441.

Horta, C.J.G. (2012) Idade Declarada nos Censos Demográficos e a Qualidade da Informação [The reported age in censuses and data quality]. In: Encontro Nacional de Estudos Populacionais, 2012, Águas de Lindóia/SP – Brasil. Anais. Belo Horizonte: ABEP, 2012.

Human Mortality Database (2019). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on 1st of May 2019).

HUMAN MORTALITY DATABASE, 2017, University of California, Berkeley (USA),, Max Planck Institute for Demographic Research (Germany), Available at www.mortality.org.br or www.humanmortality.de.

IBGE (2003). Censo demografico 2000-resultados da pesquisa de avaliacao da cobertura da coleta. Technical report, Instituto Brasileiro de Geografi a e Estatıstica - Ministerio do Planejamento, Orcamento e Gestao.

IBGE (2004). Tendencias demograficas : Uma analise dos resultados da amostra do censo demografico 2000. Technical report, Instituto Brasileiro de Geografia e Estatıstica - Ministerio do Planejamento, Orcamento e Gestao.

Instituto Brasileiro de Geografia e Esstatistíca (1960). Censo Demográfico de 1960 [The 1960 Brazilian Census]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Esstatistíca (2001). Censo Demográfico 2000: características da população e dos domicílios: resultados do universo [The 2000 Brazilian Census: population and housing charactersitcs]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (1950). Recenseamento Geral do Brazil 1 de setembro de 1940 [Brazilian census conducted in the 1st of September 1940]. Rio de Janeiro. Vol 2.

Instituto Brasileiro de Geografia e Estatistíca (1956). Brasil Censo Demográfico de 1950 [The 1950 Brazilian Census]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (1973). Censo Demográfico de 1970 [The 1970 Brazilian Census]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (1981). Sinopse preliminar do Censo Demográfico de 1980 [Premilimary results of the 1980 Brazilian Census]. Rio de Janeiro Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (1994). Censo Demográfico 1991 [The 1991 Brazilian Census]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (2010). Observações sobre a evolução da mortalidade no Brasil: o passado, o presente e perspectivas [The evolution of mortality in Brazil, past, present and perspectives]. Rio de Janeiro. Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (2011) Censo Demográfico 2010: características da população e dos domicílios: resultados do universo [The 2010 Brazilian Census: population and housing charactersitcs]. Rio de Janeiro Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (2011) Censo Demográfico 2010: características da população e dos domicílios: resultados do universo [The 2010 Brazilian Census: population and housing charactersitcs]. Rio de Janeiro Vol 1.

Instituto Brasileiro de Geografia e Estatistíca (2013). Projeção da população do Brasil por sexo e idade para o período 2000/2060 [Population projection by age and sex in Brazil from 2000 to 2060]. Rio de Janeiro.

IPUMS-I (2019). Integrated public use microdata series, international: Version 7.2[dataset]. Minesota Population Center, Minneapolis.

JASILIONIS D. , RADNÓTI L., 2016, About mortality data for Hungary, The Human Mortality Database. Background and documentation.

JASILIONIS D., 2017, About mortality data for Estonia, The Human Mortality Database. Background and documentation.

JASILIONIS D., 2017, About mortality data for Latvia, The Human Mortality Database. Background and documentation.

JASILIONIS D., 2017, About mortality data for Lithuania, The Human Mortality Database. Background and documentation.

JASILIONIS D., MESLÉ F., SHKOLNIKOV V., VALLIN J., 2011, Recent life expectancy divergence in Baltic countries, European Journal of Population 27 (4), 403–431.

Jdanov, D. A., Jasilionis, D., Soroko, E. L., Rau, R., & Vaupel, J. W. (2008). Beyond the Kannisto-Thatcher Database on Old Age Mortality: an assessment of data quality at advanced ages. IDEAS Working Paper Series from RePEc, 49(0).

Johnson-Greene, D., Dehring, M., Adams, K. M., Miller, T., Arora, S., Beylin, A., and Brandon, R. (1997). Accuracy of self-reported educational attainment among diverse patient populations: a preliminary investigation. Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists,12:635–643.

Kane, T., Rouse, C. E., and Staiger, D. (1999). Estimating returns to schooling when schooling is misreported. Technical report, NBER Working Paper No. 7235.

Kannisto, V., Lauritsen, J., Thatcher, A. R., & Vaupel, J. W. (1994). Reductions in Mortality at Advanced Ages: Several Decades of Evidence from 27 Countries. Population and Development Review, 20(4), 793-810.

KC, S., Barakatand, B., Goujon, A., Skirbekk, V., and Lutz, W. (2010). Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. Demographic Research, 22:383–472.

Kestenbaum, B. M., & Ferguson, B. R. (2005). Number of Centenarians in the United States 01/01/1990, 01/01/2000, and 01/01/2010 Based on Improved Medicare Data. Paper presented at Population Association of America, Philadelphia, Pennsylvania, March 31 – April 2 2005.

Koch, E., Romero, T., Manrıquez, L., Paredes, M., Ortuzar, E., Taylor, A., Roman, C., Kirschbaum, A., and Dıaz, C. (2007). Desigualdad educacional y socioecon´omica como determinante de mortalidad en chile: analisis de sobrevida en la cohorte del proyecto san francisco. Revista medica de Chile, 135(11).

Leeson, G.W. (2017) The impact of mortality development on the number of centenarians in England and Wales. J Pop Research 34(1), 1-15

LEON D., 2011, Trends in European life expectancy: A salutary view, International Journal of Epidemiology 40 (2), 271–277.

LEON D., CHENETL., SHKOLNIKOV V., ZAKHAROV S., SHAPIRO J., RAKHMANOVAG., VASSINS., MCKEEM., 1997, Huge variation in Russian mortality rates 1984–94: Artefact, alcohol, or what?, The Lancet 350 (9075), 383–388.

Lerch, M., Spoerri, A., Jasilionis, D., and Fernandez, F. V. (2017). On the plausibility of socioeconomic mortality estimates derived from linked data: a demographic approach. Population Health Metrics, 15(1).

Lewis, J. R. (1986). International labour migration and uneven regional development in labour exporting countries. Tijdschrift voor Economische en Sociale Geografie, 77(1):27–41.

Lima, E. E. C. and Queiroz, B. L. (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 Saude Publica, 30(8):1721–1730.

Lutz, W. and KC, S. (2011). Global human capital: Integrating education and population. Science, 333(6042):587–592.

Lutz, W. and Kebede, E. (2018). Education and health: Redrawing the preston curve. Population and Development Review, 44(2):343–361.

Lutz, W., Butz, W. P., and KC, S. (2014). World Population and Human Capital in the Twenty-First Century. Oxford University Press.

Mackenbach, J. P., Kunst, A. E., Groenhof, F., Borgan, J. K., Costa, G., Faggiano, F., Jozan, P., Leinsalu, M.,Mar tikainen, P., Rychtarikova, J., and Valkonen, T. (1999). Socioeconomic inequalities in mortality among women and among men: an international study. American Journal of Public Health, 89(12):1800–1806.

MACKENBACK J., 2013, Convergence and divergence of life expectancy in Europe: A centennial view, European Journal of Epidemiology 28 (3), 229–240.

Medford, A., Christensen, K., Skytthe, A., & Vaupel, J. W. (2019). A Cohort Comparison of Lifespan After Age 100 in Denmark and Sweden: Are Only the Oldest Getting Older? Demography, 56(2), 665–677.

MESLÉ F. , 1996, Mortality in Eastern and Western Europe: A widening gap, in Coleman D. (ed.), Europe's population in the 1990s, New York, Oxford University Press, 127–43.

MESLÉ F. , 2004, Mortality in Central and Eastern Europe: Long-term trends and recent upturns, Demographic Research, Special Edition 2 (3), 45–70.

MESLÉ F. , VALLIN J. ., 2002, Mortality in Europe: The Divergence between East and West, Population (English Edition, 2002) 57 (1), 157–197.

MESLÉ F. , VALLIN J., 2017, The end of East–Westdivergence in European life expectancies? An introduction to the special issue, European Journal of Epidemiology 33 1–13.

MESLÉ F., VALLIN J., PYROZHKOV S. ., 2012, Impact of major groups of causes on life expectancy trends, in MESLÉ, F. , VALLIN, J. (eds.), Mortality and causes of death in 20th century Ukraine, Demographic Research Monographs 9, Dordrecht, Springer, 173–195.

MÉSZÁROS J. , JASILIONIS D., 2015, About mortality data for Slovak Republic. The Human Mortality Database. Background and documentation.

Mortara, G. (1949). Comparação entre os números dos brasileiros natos, apurados pelos censos de 1940 e de 1920, e os calculados conforme determinadas hipóteses acêrca da população, da natalidade e da mortalidade [Comparison between the number of native-born Brazilians in the 1940 and 1920 census and the numbers derived from assumptions regarding the population, births and mortality]. Revista Brasileira de Estatística, 39.

MUSTARD F., 1996, East West life expectancy gap and socio-economic change, in Hertzman C., Shona K., Bobak M. (eds.), East–Westlife expectancy gap in Europe: Environmental and non-environmental determinants, Dordrecht, Springer Netherlands, 3–15.

Myers, R. J. (1966). Validity of Centenarian Data in the 1960 Census. Demography, 3(2), 470-476.

Nepomuceno, M. R. and Turra, C. M. (2019). The population of centenarians in brazil: historical estimates from 1900 to 2000. Technical report, MPIDR Working Paper WP-2019-015.

NOLTE E., SHKOLNIKOV V. , MCKEE M., 2000, Changing mortality patterns in East and West Germany and Poland. II: Short-term trends during transition and in the 1990s, Journal of Epidemiology & Community Health 54, 899–906.

OMRAN A., 1971, The epidemiologic Transition: A Theory of the Epidemiology of Population Change, The Milbank Memorial Fund Quarterly, 49, 509-538.

Paes, N. A. (2007). Qualidade das estatısticas de obitos por causas desconhecidas dos estados brasileiros. Revista de Saude Publica, 41(3):436–445.

Paes, N. A. and Albuquerque, M. E. E. (1999). Avaliacao da qualidade dos dados populacionais e cobertura dos registros de obitos para as regioes brasileiras. Revista de Saude Publica, 33(1):33–43.

Poon, L. W., & Cheung, S. L. K. (2012). Centenarian research in the past two decades. Asian Journal of Gerontology & Geriatrics, 7(1), 8–13.

Potter, J. E., Schmertmann, C. P., Assuncao, R. M., and Cavenaghi, S. M. (2010). Mapping the timing, pace, and scale of the fertility transition in Brazil. Population and Development Review, 36(2):283–307.

Preston, S. and Taubman, P. (1994). Demography of Aging, chapter Socioeconomic differences in adult mortality and health status, pages 279–318. National Academy Press.

Preston, S. H. and Coale, A. J. (1982)."Age Structure, Growth, Attrition, and Accession: A New Synthesis". Population Index 48 (2): 217-259.

Preston, S. H., Elo, I. T., & Stewart, Q. (1999). Effects of age misreporting on mortality estimates at older ages. Population Studies, 53(2), 165–177.

Preston, S. H., Heuveline, P., Guillot M. (2001). "Demography: measuring and modeling population processes". Oxford: Blackwell Publishers Ltd

PYROZHKOV S., FOYGT N., JDANOV D., 2015, About mortality data for Ukraine. The Human Mortality Database. Background and documentation.

Rau, R., Soroko, E., Jasilionis, D., & Vaupel, J. W. (2008). Continued reductions in mortality at advanced ages. Population and Development Review, 34(4), 747–768.

Renterıa, E. and Turra, C. M. (2009). Measuring educational differences in mortality among women living in highly unequal societies with defective data: the case of Brazil. Technical Report 348, Textos para Discussao Cedeplar-UFMG - Universidade Federal de Minas Gerais.

Rigotti, J. I. R. and Cerqueira, C. A. (2004). Introducao a demografia da educacao, chapter Variaveis de educacao dos censos demograficos brasileiros de 1960 a 2000, pages 71–88. Associacao Brasileira de Estudos Populacionais.

Rios-Neto, E. L. G., de Meireles Guimaraes, R. R., Pimenta, P. S. F., and de Azevedo Moraes, T. (2010). Analise da evolucao de indicadores educacionais no brasil:1981 a 2008. Technical report, Texto para discussao- Universidade Federal de Minas Gerais.

Rios-Neto, E. L. G., Miranda-Ribeiro, A., and Miranda-Ribeiro, P. (2018). Fertility differentials by education in Brazil: From the conclusion of fertility to the onset of postponement transition. Population and Development Review, 44(3):489–517.

Robine, J. M., & Paccaud, F. (2005). Nonagenarians and centenarians in Switzerland, 1860-2001: A demographic analysis. Journal of Epidemiology and Community Health, 59(1), 31–37.

Robine, J. M., Cheung, S. L. K., Saito, Y., Jeune, B., Parker, M. G., & Herrmann, F. R. (2010). Centenarians today: New insights on selection from the 5-COOP study. Current Gerontology and Geriatrics Research, 2010(120354), 1-9.

Robine, J.-M., & Vaupel, J. W. (2001). Supercentenarians: slower ageing individuals or senile elderly? Experimental Gerontology, 36(4–6), 915–930.

Robine, Jean-Marie, & Cubaynes, S. (2017). Worldwide demography of centenarians. Mechanisms of Ageing and Development, 165(B), 59-67.

Robine, Jean-Marie, & Saito, Y. (2003). Survival Beyond Age 100 : The Case of Japan Survival Beyond Age 100 : The Case of Japan. Population and Development Review, 29(supplement), 208–228.

Robine, Jean-Marie, and Vaupel, J.W. 2001. “Supercentenarians: Slower Ageing Individuals or Senile Elderly?” Experimental Gerontology 36 (4–6): 915–30.

Robine, Jean-Marie, Saito, Y., & Jagger, C. (2009). The relationship between longevity and healthy life expectancy. Quality in Ageing and Older Adults, 10(2), 5–14.

Rosenwaike, I. (1979). A New Evaluation of United States Census Data on the Extreme Aged. Demography, 16(2), 279–288.

Rosenwaike, I., & Logue, B. (1983). Accuracy of Death Certificate Ages for the Extreme Aged. Demography, 20(4), 569–585.

Rosero-Bixby, L. and Dow, W. H. (2009). Surprising SES gradients in mortality, health, and biomarkers in a latin american population of adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 64B(1):105–117.

Rostron, B., Boies, J., and Arias, E. (2010). Education reporting and classification on death certificates in the united states. Technical report, National Center for Health Statistics Vital and Health Statistics.

RYCHTAŘIKOVÁ J., 2014, The case of the Czech Republic: Determinants of the Recent Favourable Turnover in Mortality, Demographic Research, Special Collection 2 (5), 105–138.

SHKOLNIKOV V. , JDANOV D., 2016, About mortality data for Russia, The Human Mortality Database. Background and documentation.

SHKOLNIKOV V. , NEMTSOV A., 1997, The anti-alcohol campaign and variation in Russian mortality, in Bobadilla J. , Costello C. (eds.), Premature death in the new independent states, Washington DC, National Academy Press, 239–261.

SHKOLNIKOV V., 2004, Introduction to the special collection of papers on "determinants of diverging trends in mortality", Demographic Research, Special Collection 2, 1–10.

SHKOLNIKOV V., 2012, Health crises and cohort mortality, in Meslé F. , VallinJ. (eds.), Mortality and causes of death in 20th-century Ukraine, Demographic Research Monographs 9, Dordrecht, Springer, 109–117.

SHKOLNIKOV V., ANDREEV E., LEON D., MCKEE M., MESLÉ F., VALLIN J., 2004, Mortality reversal in Russia: The story so far, Hygiea Internationalis 4 (4), 29–80.

SHKOLNIKOV V., ANDREEV E., MCKEE M., LEON D., 2013, Components and possible determinants of the decrease in Russian mortality in 2004–2010, Demographic Research 28 (32), 917–950.

SHKOLNIKOV V., MESLÉ F., VALLIN J., 1997, Recent trends in life expectancy and causes of death in Russia, 1970–1993, in Bobadilla J. , Costello C. (eds.), Premature death in the new independent states, Washington DC, National Academy Press, 34–65.

Shkolnikov, V. M., Jasilionis, D., Andreev, E. M., Jdanov, D. A., Stankuniene, V., and Ambrozaitiene, D. (2007). Linked versus unlinked estimates of mortality and length of life by education and marital status: Evidence from the first record linkage study in lithuania. Social Science & Medicine, 64(7):1392–1406.

Siegel, J.S. and Passel J. S. (1976). New Estimates of the Number of Centenarians in the United States. Journal of the American Statistical Association 71(355), 559-566.

Singh, S. and Casterline, J. (1985). Reproductive change in developing countries: insights from the World Fertility Survey, chapter The socio-economic determinants of fertility, pages 199–222. Oxford University Press.

Skytthe, Axel and Jeune, Bernard. 1995. Danish centenarians after 1800. In: B. Jeune and J. Vaupel (Eds). Monographs on Population Aging. 2:55-66. Odense: Odense University Press.

Smith-Greenaway, E. and Yeatman, S. (2019). Unrealized educational expectations and mental health: Evidence from a low-income country. Social Forces.

Sorlie, P. D. and Johnson, N. J. (1996). Validity of education information on the death certificate. Epidemiology, 7(4):437–439.

Thatcher, A. R. (1992). Trends in numbers and mortality at high ages in England and Wales. Population Studies 46 (3): 411-426.

Thatcher, A. R. (2001). The demography of centenarians in England and Wales. Population: An English Selection 13(1): 139-156.

Turra, C. M. (2012). Os limites do corpo: A longevidade em uma perspectiva demografica. Revista da UFMG, 19(1):156–181.

Turra, C. M., Renteria, E., and Guimaraes, R. (2016). The effect of changes in educational composition on adult female mortality in Brazil. Research on Aging, 38(3):283–298.

Turra, C. M., Ribeiro, M. M., and de Xavier Pinto, C. C. (2018). Padroes de mortalidade por escolaridade no brasil: evidencias a partir do sistema de informacao sobre mortalidade. In Anais do XXI Encontro Nacional de Estudos Populacionais.

United Nations (2015). The millennium development goals report. Technical report, United Nations.

United Nations (2016). The sustainable development goals report. Technical report, United Nations.

United Nations (2017). World population prospects 2017. Department of Economic and Social Affairs, Population Division.

United Nations, Department of Economic and Social Affairs, Population Division (2017). World Population Prospects 2017, Online Edition.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019, Online Edition.

VALLIN J., MESLÉF., 2004, Convergences and divergences in mortality, Demographic Research, Special Collection 2, 11–44.

Vaupel, J. W., & Gowan, A. E. (1986). Passage to Methuselah: some demographic consequences of continued progress against mortality. American Journal of Public Health, 76(4), 430–433.

Vaupel, J.W., Carey, J. R., Christensen, K., Johnson, T.E., Yashin, A. I. et al. (1998). Biodemographic trajectories of longevity. Science 280(5365) 855-60.

VELKOVA A., WOLLESWINKEL-VAN DEN BOSCH J. , MACKENBACH J., 1997, The East–West life expectancy gap: Differences in mortality from conditions amenable to medical intervention, International Journal of Epidemiology 26 (1), 75–84.

VOGT T. , KLUGE F., 2013, Can public spending reduce mortality disparities? Findings from East Germany after reunification, The Journal of the Economics of Ageing 5, 7–13.

VOGT T., 2013, How Many Years of life did the fall of the Berlin Wall add? A projection of East German life expectancy, Gerontology 59 (3), 276–282.

VOGT T., VAN RAALTE A., GRIGORIEV P., MYRSKYLÄ M., 2017, The German East–West mortality difference: Two crossovers driven by smoking, Demography 54 (3), 1021–1071.

VOGTT. , MISSOV T., 2017, Estimating the contribution of mortality selection to the East-West German mortality convergence, Population Health Metrics 15 (33), 1–8.

Wang, Z., Zeng, Y., Jeune, B., & Vaupel, J. W. (1998). Age validation of Han Chinese centenarians. Genus, 54(1–2), 123–141.

WATSON P., 1995, Explaining rising mortality among men in Eastern Europe, Social Science & Medicine 41 (7), 923–934.

Wilmoth, J. (1995). The Earliest Centenarians: a statistical analysis. In: Jeune, B., Vaupel, J.W. (Eds.). Exceptional Longevity: From Prehistory to the Present. Odense University Press, Odense, 125-169.

Wilmoth, J. and Lundström, H. (1996). Extreme Longevity in Five Countries Presenteation of Trends with Special Attention to Issues of Data Quality. European Journal of Population 12, 63-93.

WORLD HEALTH ORGANIZATION, 1995, Health in Europe, Copenhagen, WHO Regional Office for Europe.

Zhu, H. and Xie, Y. (2007). Socioeconomic differentials in mortality among the oldest old in china. Research on Aging, 29(2):125–143.