Linking Birth Registration Delays to Census Age Misreporting in Brazil: Historical Patterns and Regional Disparities*

Laura Carvalho Andrade
Cássio Maldonado Turra

Abstract: Age misreporting in demographic data poses significant challenges for demographic analysis, particularly in low- and middle-income countries. This study examines two interrelated issues: (i) age data quality in Brazilian censuses (1970–2022) and (ii) its relationship to delayed birth registration. We evaluated age reporting quality using Whipple and Myers indices, age ratios, and intercensal survival ratios, focusing on the older population (aged 80+). Results reveal substantial improvements in digit preference, yet systematic age exaggeration persists, particularly in the North and Northeast regions. Additionally, we analyzed previously unpublished state-level birth registration data (1974–2021) through cohort and period approaches to examine temporal and geographic patterns of registration delays, documenting a significant decline in delayed registrations since the 1970s, albeit with persistent regional disparities. Strong positive correlations between delayed birth registration prevalence and age misreporting indicators demonstrate that historical registration deficiencies continue to undermine age data reliability. To our knowledge, this is the first study to systematically link birth registration quality to census age-data accuracy in Brazil. Our findings have broader implications for low- and middle- income countries with similar registration challenges, suggesting that observed patterns among older adults may reflect administrative data artifacts rather than genuine demographic phenomena.

1. Background

Population aging began first in high-income countries. In most Latin American and Caribbean (LAC) and Asian countries, it became more evident only after the 1970s. Since then, the process has progressed rapidly in these regions. In LAC, the proportion of individuals aged 80 and older increased from 0.41% in 1970 to 1.73% in 2020 and is projected to reach about 5% by 2050. Brazil follows a similar trend: in 1970, only 0.28% of the population was aged 80 or older, rising to 1.75% in 2020 and expected to reach about 6% by 2050 (World Population Prospects, 2024)

The demographic aging of populations is primarily driven by a sustained decline in fertility rather than by reductions in mortality rates. As Coale (1964) noted, the global decline in death rates in the early 20th century initially resulted in a younger population. This occurred because "typical improvements in health and medicine produce the greatest increases in survivorship among the young, rather than the old" (Coale, 1964, p. 50). However, more recently, survival gains have

*This paper is based on the author's master's thesis defended at the Center for Development and Regional Planning (Cedeplar), Federal University of Minas Gerais (UFMG), Brazil, under the supervision of Professor Cássio Maldonado Turra.

increasingly been concentrated among the population's oldest groups, further accelerating the aging process (Lundquist, Anderton, and Yaukey, 2014).

The increasing survival rates among older adults have sparked a growing interest in the study of adult mortality (Rau et al., 2008; Grushka, 1996). However, this interest is often limited by challenges related to data availability and reliability. Some scholars have focused on exploring the potential upper limits of human life expectancy (Olshansky, Carnes, and Cassel, 1990; Oeppen and Vaupel, 2002) and the existence of mortality plateaus at extremely advanced ages (Rau et al., 2017; Barbi et al., 2018). In contrast, others argue that estimates at very old ages are severely compromised by data inaccuracies, such as age misreporting in both population and death counts (Gavrilov and Gavrilova, 2019; Newman, 2024).

Inaccurate data on age, particularly in low- and middle-income countries, poses significant challenges (Bhat 1990). In Brazil, uncertainty about the reliability of age data for older individuals has been one reason why, until 2021, the Brazilian Census Bureau published complete life tables that only extended to 80+ years, rather than to 110+ years. In contrast, countries with high-quality demographic data and initiatives such as the Human Mortality Database (HMD, 2025) regularly produce life tables that extend to very advanced ages, often up to 110 years or more.

There are two common types of errors related to age reporting: age heaping and systematic age misreporting (Siegel and Swanson, 2004; Palloni, Beltrán-Sánchez, and Pinto, 2021). Age heaping, also known as digit preference, occurs when individuals round their reported ages, often favoring numbers ending in 0 or 5, even when more precise information is requested (Myers, 1966). According to Siegel and Swanson (2004), age heaping is the most frequent type of error found in data recorded by single years of age, particularly among populations with low levels of education. Fortunately, age heaping is relatively easy to detect and correct (Palloni, Beltrán-Sánchez and Pinto, 2021), and it can be significantly mitigated by grouping data into broader age categories, such as five-year age intervals.

On the other hand, systematic misreporting of age introduces a more complex challenge, as it is harder to diagnose and adjust (Palloni, Beltrán-Sánchez and Pinto, 2021). This type of error can introduce bias into demographic indicators, including mortality rates (Coale and Kisker, 1986; Dechter and Preston, 1991; Preston et al., 1996; Preston, Elo and Stewart, 1999; Palloni, Beltrán-Sánchez and Pinto, 2021).

Ewbank (1981) argues that addressing age misreporting requires understanding the underlying sources of these reporting errors. According to the author, such errors often arise from respondents' limited knowledge of their own exact age or birth date, as well as those of other household members in cases of proxy reporting. Ewbank also emphasizes that errors in recorded ages are shaped by the methods and procedures used during data collection and processing. Factors such as interviewer bias and data handling issues, including miscoding, keypunching mistakes, or the statistical imputation of missing values, can further exacerbate age misreporting.

In an effort to minimize age misstatement, various household surveys have included questions about both declared age and date of birth, later reconciling the responses to ensure consistency. This strategy has shown improved results in countries such as England and Wales and the United States, primarily by reducing severe age heaping (Ewbank, 1981, p. 12). Nonetheless, this approach can give rise to two potential issues. Firstly, Myers (1976) demonstrated that asking about the date of birth may lead to what he calls "reverse heaping" (a tendency for people to report birth years ending in zeroes or fives). This, in turn, produces age heaping at different digits depending on the reference year. Second, individuals, especially those from lower socioeconomic backgrounds, may lack accurate knowledge of their actual birth dates, particularly in countries where birth registration systems are weak or incomplete (Ewbank, 1981; Turra et al., 2023).

Research findings have consistently demonstrated the prevalence of age misreporting across a wide range of populations (Ortega and Garcia, 1985; Rosenwaike and Preston, 1984; Coale and Kisker, 1986; Dechter and Preston, 1991; Grushka, 1996; Preston et al., 1996; Richman, 2017). Moreover, the quality of age reporting tends to worsen with advancing age, largely due to an "almost universal" tendency to exaggerate reported ages (Coale and Kisker, 1986, p. 401). A clear indication of this exaggeration appears in centenarian data. Myers (1966) identified substantial overstatement in U.S. censuses, noting that while the 1960 census recorded over 10,000 centenarians, his estimates pointed to a more realistic figure of around 3,700. Similar patterns have been documented in Brazil. For example, the 1991 Brazilian census reported over 13,000 centenarians, but Gomes and Turra (2009) estimated the number to be closer to 4,500 using the extinct generations method. Likewise, Nepomuceno and Turra (2020) documented persistent overstatement of centenarians across multiple Brazilian censuses throughout the twentieth century.

This tendency toward age exaggeration has important implications, as the systematic overstatement of age can result in the underestimation of mortality rates among older age groups (Coale and Kisker, 1986; Dechter and Preston, 1991; Grushka, 1996; Preston, Elo and Stewart, 1999; Palloni, Beltrán-Sánchez and Pinto, 2021). This bias may lead to so-called "mortality crossovers", a phenomenon in which, as Coale and Kisker (1986, p. 389) explain, "the death rates in the less favored population, higher early in life, cross the rates in the other population, and are lower in the later years."

In the Brazilian context, demographic data are not immune to age misreporting (Turra et al., 2023). In addition to the well-documented over-enumeration of centenarians in national censuses (Gomes and Turra, 2009; Reis and Turra, 2016; Nepomuceno and Turra, 2020), further evidence suggests the effect of systematic age misreporting on Brazilian mortality data. For instance, Turra (2012) analyzed female mortality rates above age 65 in 2009 Brazil and compared them to those of France, Japan, and Sweden in earlier periods (when these countries had a similar life expectancy at age 45 to that of Brazil in 2009). The analysis revealed that Brazil's mortality curve flattened at older ages, resulting in a mortality crossover, where Brazilian mortality rates fell below those of Sweden, Japan, and France. These results are particularly concerning in light of Coale and Kisker's (1986) observation that such crossovers frequently arise in contexts marked by significant age misreporting, leading to an overestimation of mortality at younger ages and an underestimation at the oldest ages.

Furthermore, Turra et al. (2023) found that requiring proof of age through official documentation improved the quality of age data among the older population by reducing both age heaping and age exaggeration. Nevertheless, inconsistencies remained. The authors proposed that these discrepancies may stem from inaccuracies in the identity cards, which were often based on erroneous birth records (a consequence of Brazil's historically weak and uneven civil registration system).

The importance of birth registration

The absence of a birth certificate can disadvantage children both early and later in life. In early childhood, studies have shown a strong association between birth registration and access to essential health services, including full immunization and vitamin supplementation (UNICEF, 2005). In a study of birth registration across Latin America and the Caribbean, Brito, Corbacho, and Osorio (2013) found that unregistered children were more likely to have limited access to vaccines and lower educational attainment. As children grow older, official documentation becomes a key protective factor against trafficking, early marriage, child labor, and prosecution as an adult when charged with a crime (UNICEF, 2005; 2013; Jayaraman et al., 2016; Hunter and Brill, 2016).

While birth registration benefits individuals, universal registration also benefits states. An efficient civil registration and vital statistics (CRVS) system, ensuring complete birth records and the registration of other demographic events, enables accurate population monitoring and reliable demographic data. This, in turn, supports effective public service planning, including school funding and healthcare infrastructure, and improves the quality of demographic indicators such as mortality schedules (UNICEF, 2005; AbouZahr et al., 2015; Hunter and Brill, 2016).

Nonetheless, the under-registration of births remains a significant global issue. As of 2012, UNICEF (2013) estimated that 230 million children under five had not had their births officially recorded. Moreover, Dunning, Gelb, and Raghavan (2014) reported that an additional 70 million had been registered but had not received a birth certificate. The problem extends into adulthood: approximately 1 billion adults worldwide lack official identity documents (Bhatia et al., 2019).

Despite these alarming numbers, significant progress has been made in birth registration coverage worldwide. Drawing on UNICEF data, Dunning, Gelb, and Raghavan (2014) analyzed trends in birth registration from 2000 to 2020 (projected estimates), showing a global decline in the proportion of unregistered children from approximately 42% in 2000 to a projected 29.2% in 2020. Unsurprisingly, global averages mask substantial regional disparities. In Central and Eastern Europe, the share of unregistered children dropped from 8% in 2000 to a projected 0.5% in 2020, while in Latin America and the Caribbean, it declined from 17% to 3.8% over the same period. Meanwhile, South Asia continued to report the highest levels, declining from 69% in 2000 to a projected 52.2% in 2020. These trends indicate that while some regions are approaching universal birth registration, others remain significantly behind (Dunning, Gelb, and Raghavan, 2014). There are also important contrasts within regions: in 2012, Chile, Cuba, and Uruguay had full birth registration coverage (at 100%), whereas Bolivia and Paraguay reported estimated coverage of around 76% (UNICEF, 2013).

Even as birth registration coverage improves globally (UNICEF, 2005, 2013; Dunning, Gelb, and Raghavan, 2014), simply registering a birth at some point during a person's life is insufficient. Registration must occur as soon after birth as possible, as stipulated in Article 7 of the United Nations Convention on the Rights of the Child. As early as the 1940s, Hedrich (1942) highlighted that birth certificates filed promptly after birth are more reliable, while those registered later tend to contain inaccuracies regarding birth dates and even birthplaces. Similar arguments have been made by AbouZahr et al. (2015), Hunter (2018), and Harbers (2020). As Hunter (2018) states, "Late registration is better than no registration, but it is suboptimal."

Consequently, in contexts with a high prevalence of delayed birth registration, age-related documents may be subject to bias. Such documents can be affected by inaccuracies arising from imprecise knowledge during late registration or potentially intentional falsification. This issue is particularly significant when these documents are used to access social benefits linked to older age or to obtain exemptions from military service obligations (Ewbank, 1981; Turra et al., 2023).

Brazil is a key example of a country that has historically experienced high levels of delayed birth registration and under-registration. Giraldelli and Wong (1984) analyzed delayed registration patterns in the state of São Paulo, one of Brazil's wealthiest states, and sought to link them to trends in under-registration. The authors identified a relationship between the number of delayed registrations in a given year and the overall level of under-registration in that same year, estimating that in the early 1980s, under-registration levels in the state ranged from 8% to 9%. Later, Wong and Turra (2007) showed that in 2004, only about 82% of births occurring in Brazil were registered within the same year, while more than 500,000 births registered in 2004 corresponded to individuals born in previous years.

Nonetheless, Brazil has made significant progress in reducing under-registration of births. In 1991, the under-registration level in the country was estimated at 29.2%, declining to 21.3% by 2000 (IBGE, 2002). By 2010, this figure had dropped to 6.6% (IBGE, 2011). The trend continued, and by 2022, IBGE (2024b) reported an all-time low under-registration level of just 1.31%. However, national averages once again concealed critical regional differences. In 1991, for example, while the Southeast region had the lowest under-registration level among regions at 6.3%, the North recorded an alarming 65.1% (IBGE, 2002). These differences have steadily narrowed over the years, but they persist. In 2022, under-registration in the Southeast was estimated at just 0.35%, while the North had a considerably higher number of 5.14% (IBGE, 2024b).

Regardless of recent improvements, it is evident that Brazil's birth registration systems have historically suffered from deficiencies. Turra et al. (2023) noted that such deficiencies may be associated with the age reporting errors prevalent in both population and death data, especially at older ages. Older individuals were born during a time when the Brazilian birth registration systems were much more limited, increasing the likelihood that their births went unreported or were only registered much later in life, at a point when their knowledge of their actual age may have been absent, inaccurate, or deliberately altered[†]. This, in turn, may lead to erroneous age information even on official identity documents (Turra et al., 2023).

Given this context, this study addresses two interrelated issues: (i) the quality of age data and spatial distribution of the older population across Brazilian states (1970–2022), and (ii) temporal and geographic patterns of delayed birth registration. First, we evaluate the quality of age-related data in Brazilian population censuses from 1970 to 2022, focusing specifically on the population aged 80 and above. Employing standard demographic indicators (including Whipple's and Myers'

[†] Brill (2013) found that eligibility for the BPC program (Benefício de Prestação Continuada) in Brazil increased the likelihood of older individuals obtaining birth certificates later in life, at 65 years old.

indexes, age ratios, and intercensal survival rates) we document persistent age heaping and age exaggeration across Brazilian regions. Second, we examine patterns of delayed birth registration in Brazil across both time and space. Through period and cohort analyses, we demonstrate how historical weaknesses in the civil registration system may have contributed to sustained inaccuracies in age documentation. Finally, we investigate the association between delayed registration prevalence and indicators of age misreporting.

To our knowledge, this is the first study to systematically examine the relationship between delayed birth registration prevalence and census age-data quality in Brazil. By linking civil registration deficiencies to age misreporting, we advance understanding of demographic data limitations in Brazil and, more broadly, in low- and middle-income countries where similar registration challenges persist.

2. Data

This analysis uses data from two primary sources:

Census Data: Data from Brazilian censuses conducted between 1970 and 2022. For the 1970, 1980, and 1991 censuses, data were sourced from IPUMS International. We obtained data from the 2000, 2010, and 2022 censuses through the Brazilian Census Bureau's system (SIDRA IBGE), which provides population counts tabulated by single years of age, sex, and geographic areas. The analysis tabulated population counts by single years of age, sex, Brazilian states, and macroregions (North, Northeast, Southeast, South, and Midwest).

Birth Registration Data: Novel data from the Brazilian Census Bureau (IBGE) containing information on birth registrations in Brazil categorized by the year of registration and the year of birth, disaggregated by sex and Brazilian state. The dataset spans the registration period from 1974 to 2021 and includes all registered births within that time frame, regardless of whether they were recorded promptly or delayed. The dataset is structured as a matrix, with birth years in the columns and registration years in the rows. Birth years range from 1900 to 2021, while registration years extend from 1974 to 2021.

3. Methods

Quality of Age-Related Census Data (1970–2022)

To evaluate the quality of Brazilian censuses from 1970 to 2022, we calculated several indicators:

Age heaping: We calculated the Whipple index to measure age heaping for ages 23 to 62, and Myers's blended index to capture age heaping among older ages (80 to 99). We calculated both indices separately by sex, state, and region for each of the six census years.

Age ratios: To further analyze the oldest segments of the population, we calculated age ratios for specific groups: 100+/80+, 90+/80+, and 100+/50+. These metrics can help identify potential errors in age reporting, such as age exaggeration. Our analysis compares these estimates across Brazilian states and census years and benchmarks them against countries with high-quality demographic data, including Sweden, Japan, France, and Italy.

"Older adults" prevalence: Our analysis estimates the prevalence of centenarians and individuals aged 80 years and older, calculated per 10,000 inhabitants for each state and sex. The results are presented on maps of Brazil that reflect the state boundaries applicable during each census year.

Intercensal survival ratios: While period measures are important, cohort measures offer valuable insights into the quality of age-related data. Therefore, to improve our analysis, we estimated intercensal survival ratios, disaggregated by state and sex. These ratios are based on life table concepts and represent the proportion of the population in a specific age group at one census compared to the population of the same birth cohort in the previous census. In populations considered closed to migration, the ratios primarily reflect mortality. However, they can also be influenced by differences in census coverage and age misreporting. In contrast, for open populations, the intercensal survival ratios capture the combined effect of mortality, census errors, and migration (Siegel and Swanson, 2004).

The survival ratios were calculated by dividing the population aged 100 years and older in 2022 by the population aged 50 years and older in 1972. To estimate the 1972 population, we used interpolation based on the mean annual growth rate between the 1970 and 1980 censuses.

Delayed Birth Registrations (1974–2021)

In a previous study, Giraldelli and Wong (1984) developed a methodology to estimate the completeness of birth registration in the state of São Paulo. Their approach assumed that "delayed birth registrations provide a good approximation of the number of births that were not declared in the year of occurrence and that will either be registered later or permanently omitted" (Giraldelli and Wong, 1984, p. 827). However, as previously mentioned, the objective of this study is different: we aim to examine the patterns of delayed birth registration in Brazil and how these patterns have evolved across Brazilian states and regions.

Building on the assumption made by Giraldelli and Wong (1984), a lower prevalence of delayed registrations suggests an improvement in the birth registration system. We hypothesize that a higher prevalence of timely registrations, as opposed to delayed ones, results in more reliable proof of age documentation and potentially enhances the quality of age-related population data (Turra et al., 2023; Hedrich, 1942).

To examine these patterns, we employed both cohort and period approaches:

Cohort perspective: We tracked birth registrations for the cohorts 1974, 1984, 1994, and 2004 over a period of 20 years after the births. We calculated the proportion of cohort births registered without any delay (timely registrations) and compared this to those with delayed registrations. The delays were categorized by the number of years, ranging from 1 to 20 years.

Period perspective: We calculated the total number of births registered in each year from 1974 to 2021. We then categorized these registrations based on the number of years delayed, ranging from timely registrations to those delayed by 30 years or more, and computed proportions of delayed registrations relative to the total number recorded in each year.

Finally, we examined potential associations between the prevalence of delayed birth registration and the quality of old-age census data by creating scatterplots and calculating Pearson correlation coefficients for various combinations of variables, separated by sex.

It is important to emphasize that our analysis does not aim to establish any causal relationships, as the data and methods employed do not permit causal inference. Instead, the objective is to examine whether there are correlations between the birth registration indicators and the age data quality indicators.

4. Results

Quality of Age-Related Census Data (1970–2022)

Whipple and Myers's indexes

The results show a significant improvement in the quality of age reporting in Brazil over the decades, with a steady decline in age heaping. Nationally, the Whipple index for the total population decreased from 126.3 in 1970 to approximately 104 in 2022, indicating stabilization and a low prevalence of heaping on digits ending in "0" and "5". Similarly, Myers's index for the total population declined by 75% between 1970 and 2022.

However, geographic differences persist over time, with clustering at specific ages more marked in states within the North, Northeast, and Midwest regions. In 1970, regional disparities were significant, with the national Whipple index average of 126.3 masking substantial variations ranging from 112.9 in the South to 148.1 in the Northeast. By 2022, the indices converged around 105, but differences remained, with the North at 105.1 and the South at 103.1.

Age Ratios

The 100+/80+ ratio, which measures the number of centenarians per 100 people aged 80 and older, has significantly declined over time at the national level, from 1.81 in 1980 to 0.83 in 2010, representing a 54% reduction (see Table 1). Between 2010 and 2022, the national ratio remained stable. Despite this general decline across all regions, significant geographic differences remain. In 2022, the North region had a ratio 2.8 times higher than the South region (1.42 vs 0.50). The evidence indicating age misreporting becomes even more compelling when compared to reference countries. In 2022, Brazil's 100+/80+ ratio was 1.8 times higher than Sweden's.

Table 1: Age ratios (per 100). Brazilian regions and countries with high-quality data. 1980–2022, Both sexes

Countries / Regions	**100+/50+**					**100+/80+**					**90+/80+**					
	1980	**1991**	**2000**	**2010**	**2022**	**1980**	**1991**	**2000**	**2010**	**2022**	**1970**	**1980**	**1991**	**2000**	**2010**	**2022**
Brazil	**0.07**	**0.07**	**0.09**	**0.06**	**0.07**	**1.81**	**1.31**	**1.34**	**0.83**	**0.82**	**15.8**	**14.4**	**11.2**	**14.3**	**15.3**	**17.1**
North	0.08	0.10	0.14	0.09	0.09	2.71	1.97	2.18	1.37	1.42	17.8	16.4	10.7	16.6	18.3	16.9
Northeast	0.09	0.10	0.11	0.11	0.12	2.06	1.42	1.34	1.20	1.31	16.9	17.0	11.9	15.5	18.4	18.6
Southeast	0.06	0.06	0.08	0.04	0.05	1.43	1.14	1.21	0.60	0.59	14.7	13.0	11.1	13.6	13.7	17.0
South	0.07	0.06	0.07	0.04	0.04	1.92	1.23	1.24	0.56	0.50	15.0	12.4	10.6	12.5	13.0	15.2
Midwest	0.10	0.07	0.11	0.06	0.05	3.49	1.73	2.11	0.98	0.78	19.8	17.0	10.4	14.9	15.1	15.5
Sweden	**0.01**	**0.02**	**0.03**	**0.04**	**0.07**	**0.13**	**0.16**	**0.21**	**0.31**	**0.46**	**8.3**	**10.0**	**11.4**	**14.4**	**16.7**	**17.8**
France	**0.01**	**0.02**	**0.04**	**0.07**	**0.10**	**0.13**	**0.18**	**0.36**	**0.48**	**0.68**	**8.9**	**10.4**	**12.3**	**19.4**	**13.8**	**23.3**
Italy	**0.01**	**0.01**	**0.02**	**0.05**	**0.07**	**0.08**	**0.13**	**0.23**	**0.36**	**0.42**	**7.2**	**9.1**	**10.1**	**15.7**	**13.0**	**18.7**
Japan	**0.00**	**0.01**	**0.04**	**0.10**	**0.20**	**0.07**	**0.18**	**0.41**	**0.67**	**0.96**	**7.3**	**8.6**	**9.0**	**14.9**	**16.7**	**21.1**

Source: IPUMS International, Sidra IBGE, and United Nations WPP 2024.

In contrast to Brazil's declining trend, reference countries have consistently increased their 100+/80+ ratios during the same period. This trend is expected due to the rising survival rates among older populations (Rau et al., 2008; Turra, 2012), a phenomenon also observed in Brazil (Campos and Rodrigues, 2004). In this context, the significant decline in Brazil's 100+/80+ ratio is noteworthy. It suggests considerable improvements in the quality of census data for older age groups, including a reduction in age misreporting

While these findings are compelling, it is important to approach comparisons of relative cohort sizes across different populations and time periods with caution. Age ratios are influenced not only by mortality rates but also by population composition effects, such as the varying original sizes of cohorts at birth and migration rates (Preston, Heuveline, and Guillot, 2001). Nevertheless, the significant differences between Brazil's ratios and those of countries with high-quality data raise concerns about the quality of the data (Gomes and Turra, 2009). These results suggest that age misreporting exists across all Brazilian regions and states. However, the extent of bias in age-related data varies considerably, with certain regions showing much more pronounced inaccuracies than others.

Older adults prevalence

In 1970, only four states (Sergipe, Ceará, Paraíba, and Rio Grande do Norte) had 50 or more men aged 80 or older per 10,000 inhabitants. In contrast, among women, 14 states exceeded this threshold, primarily in the Northeast and Southeast regions. By 1991, the prevalence of individuals aged 80 or older had increased significantly. For the first time, both men and women in certain states reached a concentration of at least 100 individuals per 10,000 inhabitants. Six states have achieved this milestone among women, with four located in the Northeast region. Similarly, for men, two states exceeded this figure, and both were also in the Northeast.

By the 2000 census, the prevalence of individuals aged 80 and older among women was below 100 per 10,000 inhabitants only in states located in the North and Midwest regions. The growth in prevalence for men was slower, reflecting women's higher probability of surviving to older ages. According to IBGE (2024), life expectancy at birth in Brazil in 2000 was 75.08 years for women compared to 67.32 years for men, resulting in a gap of 7.76 years. Nevertheless, the 2022 map reveals that the prevalence of individuals aged 80 and above increased for both sexes, with only a few northern states showing the lowest prevalence levels. This upward trend aligns with the population aging process that began in Brazil following the decline in fertility in the 1960s (Carvalho and Garcia, 2003).

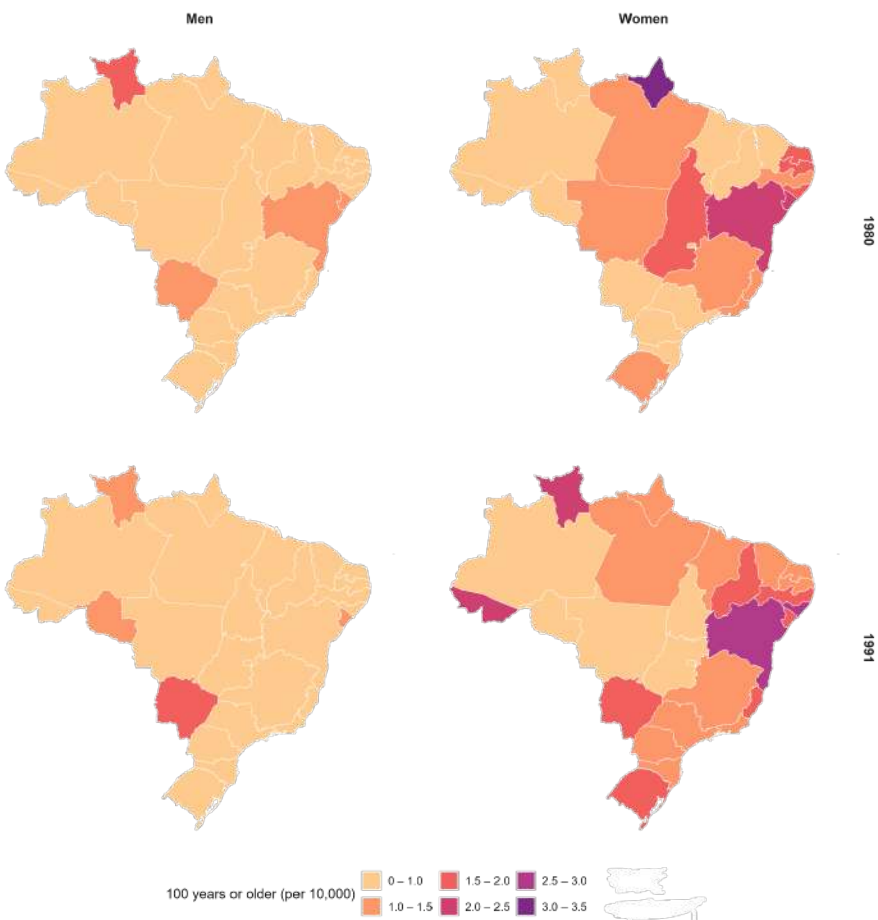
Figures 1 and 2 present the changes in the distribution of centenarians in Brazil from 1980 to 2022. During this time, the prevalence of centenarians increased significantly across all states, mirroring the trends observed in the population aged 80 years and older.

In 1980, only about three states (Bahia, Sergipe, and Amapá), all located in the Northeast and North regions, had two or more female centenarians per 10,000 female inhabitants. Among these states, Amapá had the highest prevalence at 3.08 per 10,000 women. By 2022, the proportion of states exceeding this threshold had risen to 66%. Notably, the ten states showing the highest prevalence were all situated in the North and Northeast regions. Bahia recorded the overall highest prevalence, with 5.21 centenarians per 10,000 women.

The number of male centenarians has historically been lower due to differences in mortality between sexes. In 1980, no states recorded at least two male centenarians per 10,000 inhabitants, and only 15% of states had at least one male centenarian per 10,000. However, by 2022, there was a significant change: 66% of states had at least one male centenarian per 10,000 inhabitants. Maranhão had the highest concentration, with 2.60 male centenarians per 10,000.

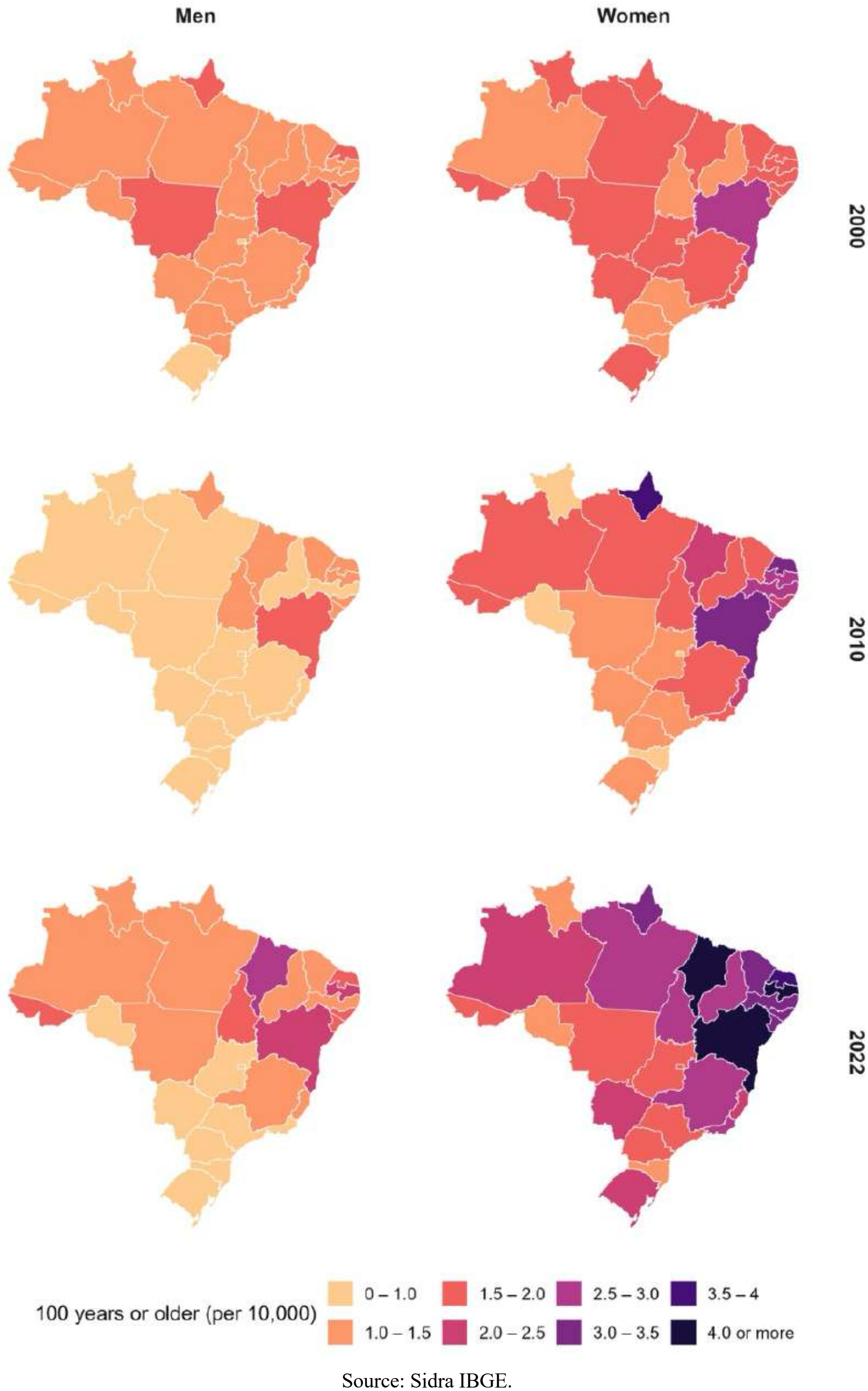
The overall increase in the prevalence of centenarians in Brazil aligns with demographic trends, as noted by Gomes and Turra (2009, p. 496): "High fertility levels in the past, combined with fast declining mortality rates at older ages have certainly allowed a much larger number of people to survive to the age 100 and older in Brazil". However, these results also indicate potential inconsistencies in the data.

Figure 1: Spatial distribution of people aged 100 years and older across Brazilian states, Men and Women – 1980, and 1991 censuses



Source: IPUMS International – Minnesota Population Center.

Figure 2: Spatial distribution of people aged 100 years and older across Brazilian states, Men and Women – 2000, 2010, and 2022 censuses



Source: Sidra IBGE.

Between 2000 and 2010, there was an important decrease in the number of individuals reporting they were 100 years or older, both in proportional and absolute terms. The total number of centenarians fell from 24,576 in 2000 to 24,236 in 2010, even as the total population grew by 12%, increasing from 169,799,170 to 190,755,799 during the same period. This decline may suggest improvements in age reporting, potentially reflecting a reduction in age exaggeration (Reis and Turra 2016; IBGE, 2000; IBGE, 2010).

The data at the state level clearly demonstrated this trend. For men, all states, except for Sergipe, which experienced a slight increase from 1.34 to 1.39 centenarians per 10,000 inhabitants, saw declines in the prevalence of centenarians from 2000 to 2010. On the other hand, among women, only 13 out of 27 states experienced decreases in the centenarian's prevalence during this period. The 14 states that showed increases were primarily located in the Southeast (Espírito Santo and Minas Gerais), North (Amapá, Amazonas, and Tocantins), and Northeast (all nine states in the region).

Finally, significant fluctuations in centenarian prevalence in certain states also raise concerns about the reliability of the data. For example, in Amapá, the estimated prevalence of female centenarians decreased from 3.08 per 10,000 inhabitants in 1980 to 1.15 in 1991. It then rose to 1.68 in 2000, increased further to 3.92 in 2010, and slightly declined to 3.30 in 2022. While some variation is expected due to the rarity of centenarians, these excessive fluctuations suggest potential inconsistencies in age reporting.

Intercensal survival ratios

Figures 3 and 4 present the intercensal survival ratios calculated using data from the 1972 interpolated population and the 2022 census. The population aged 100 years and older in 2022 was divided by the corresponding cohort 50 years earlier, represented by individuals aged 50 years and older in 1972. All calculations were performed separately by sex and state of residence at the time of the census interview.

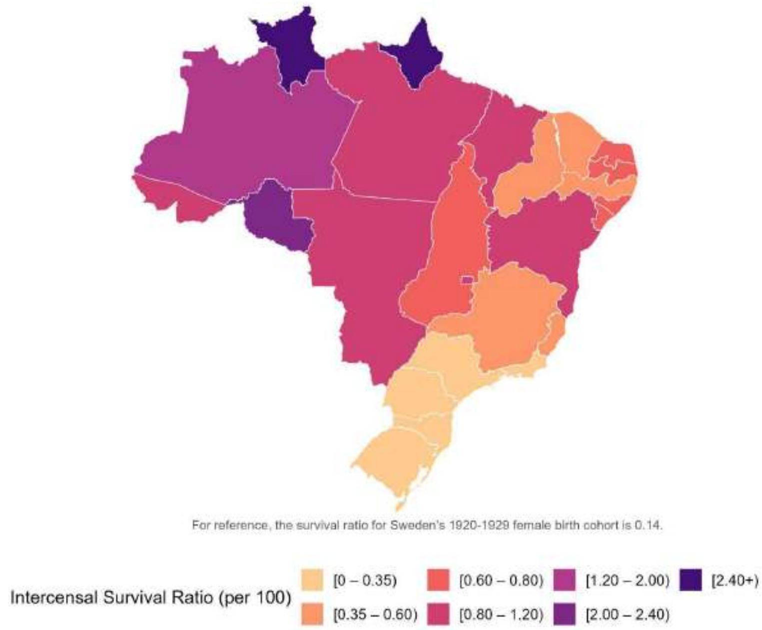
At the national level, the intercensal survival ratio was estimated at 0.49% for women and 0.20% for men. If these ratios primarily reflect mortality, this indicates that approximately 0.49% of women aged 50 years and older in 1972 survived to age 100 or older by 2022, compared to 0.20% of men from the same birth cohort.

Regional variations in survival ratios were even more pronounced in this set of calculations compared to those based on state of birth. Once again, the North region recorded the highest ratios, with 1.22% for women and 0.53% for men. The Midwest region followed, showing ratios of 0.78% for women and 0.35% for men. The Northeast region had ratios of 0.71% for women and 0.35% for men. In contrast, the Southeast and South regions reported the lowest survival ratios, with the Southeast at 0.37% for women and 0.12% for men, and the South at 0.29% and 0.10% for women and men, respectively.

Among women, only the southern states and two southeastern states, São Paulo and Rio de Janeiro, reported survival ratios below 0.35%. These values ranged from 0.27% in Rio Grande do Sul to 0.32% in both São Paulo and Rio de Janeiro. Ten states recorded survival ratios between 0.35% and 0.80%, including Minas Gerais, Espírito Santo, Goiás, and nearly all Northeastern states, except Bahia and Maranhão. The remaining states had survival ratios exceeding 0.80%. The highest survival ratios were concentrated in Northern states, where significant outliers were noted. Notably, Amapá and Roraima reported exceptionally high survival ratios of 2.78% and 2.92%, respectively.

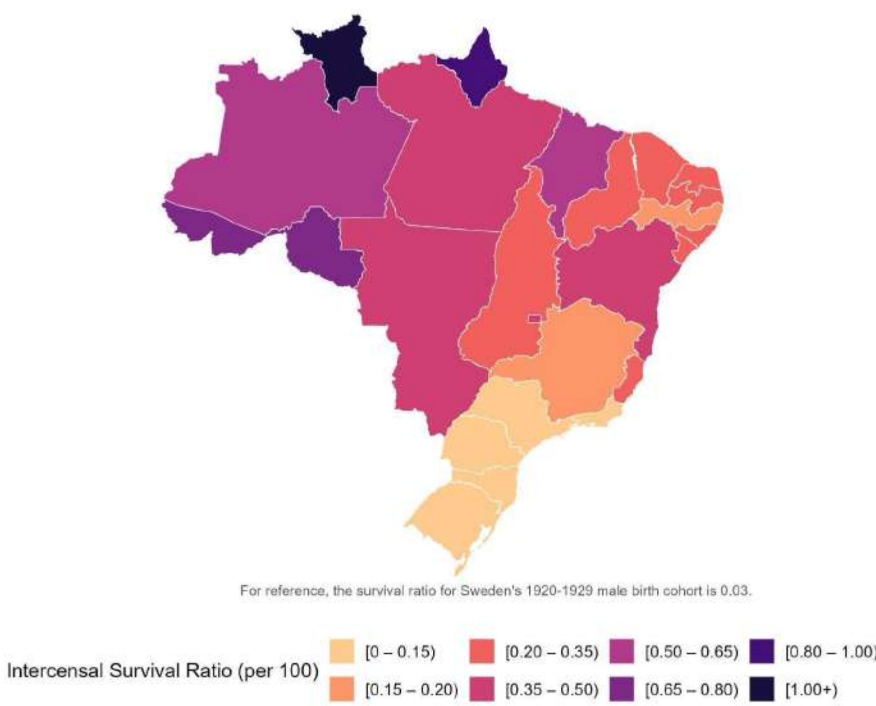
The results for men (Figure 4) were similar to those for women (Figure 3). The lowest survival ratios, falling below 0.15%, were recorded in the southern states and the southeastern states of São Paulo and Rio de Janeiro. In contrast, the highest ratios were observed in the Northern states, the Midwestern state of Mato Grosso, and the Northeastern states of Bahia and Maranhão. Just like with the women, the two states showing the highest survival ratios for men were Amapá and Roraima, with values of 0.88% and 1.65%, respectively.

Figure 3: Spatial distribution of **intercensal survival ratios for women** (%): Population aged 100+ in the 2022 census relative to the same cohort aged 50+ in 1972 (obtained through interpolation), by state of residence. Brazil.



Source: IPUMS International – Minnesota Population Center; Sidra IBGE.

Figure 4: Spatial distribution of **intercensal survival ratios for men** (%): Population aged 100+ in the 2022 census relative to the same cohort aged 50+ in 1972 (obtained through interpolation), by state of residence. Brazil.



Source: IPUMS International – Minnesota Population Center; Sidra IBGE.

The survival ratios were notably higher compared to those observed in the 1920–1929 birth cohorts in Sweden and Italy, which serve as benchmarks. In Sweden, the survival ratios of individuals aged 100 and older (to those aged 50 and older) were 0.14% for women and 0.03% for men. In Italy, these ratios were 0.18% for women and 0.05% for men (HMD, 2024).

Delayed Birth Registrations (1974–2021)

Cohort perspective

Results indicate a significant improvement in the rate of timely birth registration in Brazil over the decades. At the national level, delayed registration was highly prevalent for the 1974 birth cohort, with 43.3% of male births and 44.3% of female births registered late. For the 2004 cohort, timely registration improved significantly, with 84.4% of births (both sexes) registered in the year they occurred (see Table 2).

Table 2: Proportion of Births (%) Registered on Time and Late, by Sex and Birth Cohort (Relative to Total Registrations Within 20 Years of Birth), 1974–2004. Brazil and Regions

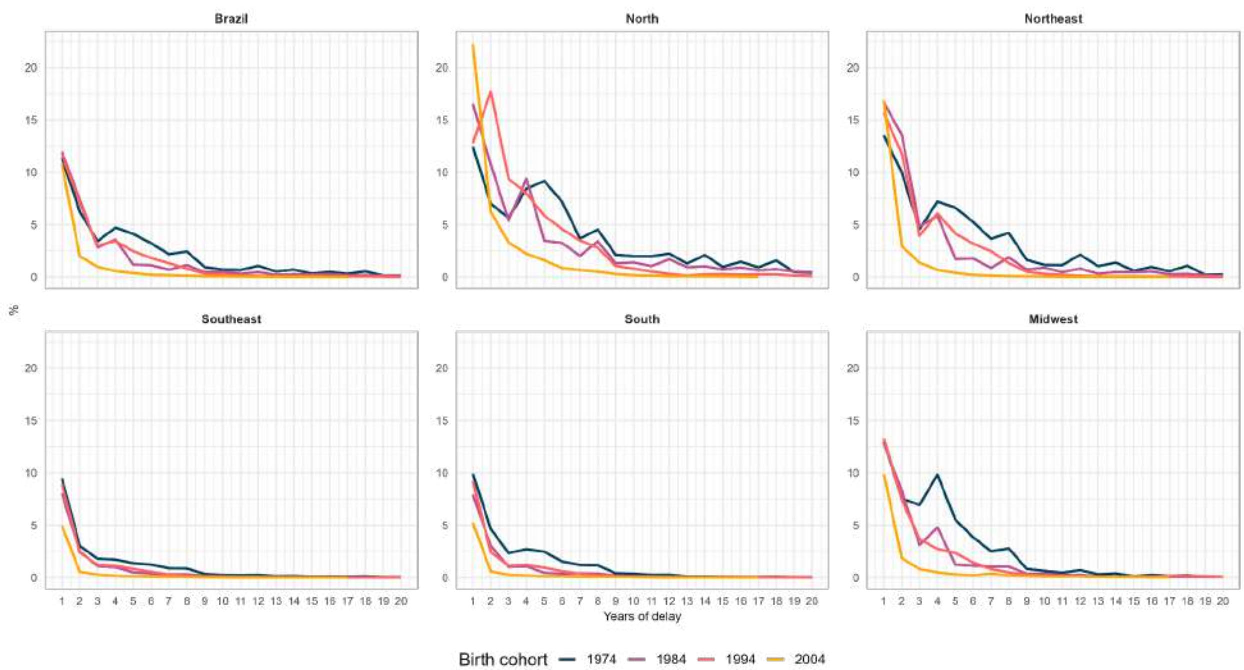
Region	Birth Cohort	1974		1984		1994		2004	
		Registered in the same year as the occurrence	Registered with a delay	Registered in the same year as the occurrence	Registered with a delay	Registered in the same year as the occurrence	Registered with a delay	Registered in the same year as the occurrence	Registered with a delay
Brazil									
	Men	56.7	43.3	67.4	32.6	67.4	32.6	84.4	15.6
	Women	55.7	44.3	66.8	33.2	67.1	32.9	84.4	15.6
North									
	Men	25.5	74.5	34.6	65.4	31.2	68.8	60.4	39.6
	Women	24.4	75.6	34.4	65.6	31.1	68.9	61.4	38.6
Northeast									
	Men	33.7	66.3	48.2	51.8	50.0	50.0	77.2	22.8
	Women	32.9	67.1	47.3	52.7	49.9	50.1	77.1	22.9
Southeast									
	Men	78.4	21.6	85.5	14.5	84.3	15.7	93.6	6.4
	Women	77.7	22.3	85.1	14.9	84.1	15.9	93.5	6.5
South									
	Men	72.8	27.2	84.9	15.1	83.4	16.6	93.0	7.0
	Women	72.1	27.9	84.5	15.5	83.3	16.7	93.0	7.0
Midwest									
	Men	44.8	55.2	65.0	35.0	66.8	33.2	85.4	14.6
	Women	43.8	56.2	64.6	35.4	66.2	33.8	85.2	14.8

Source: IBGE.

However, national averages conceal considerable regional disparities. The North and Northeast regions have significantly higher rates of delayed registration compared to the Southeast and South regions. For the 1974 cohort, the North region had 74.5% of male births and 75.6% of female births registered late, while the Southeast had only 21.6% and 22.3%, respectively. By the 2004 cohort, the Southeast and South regions recorded approximately 93% of timely birth registrations, while the Northeast and North lagged at 77% and 60%, respectively.

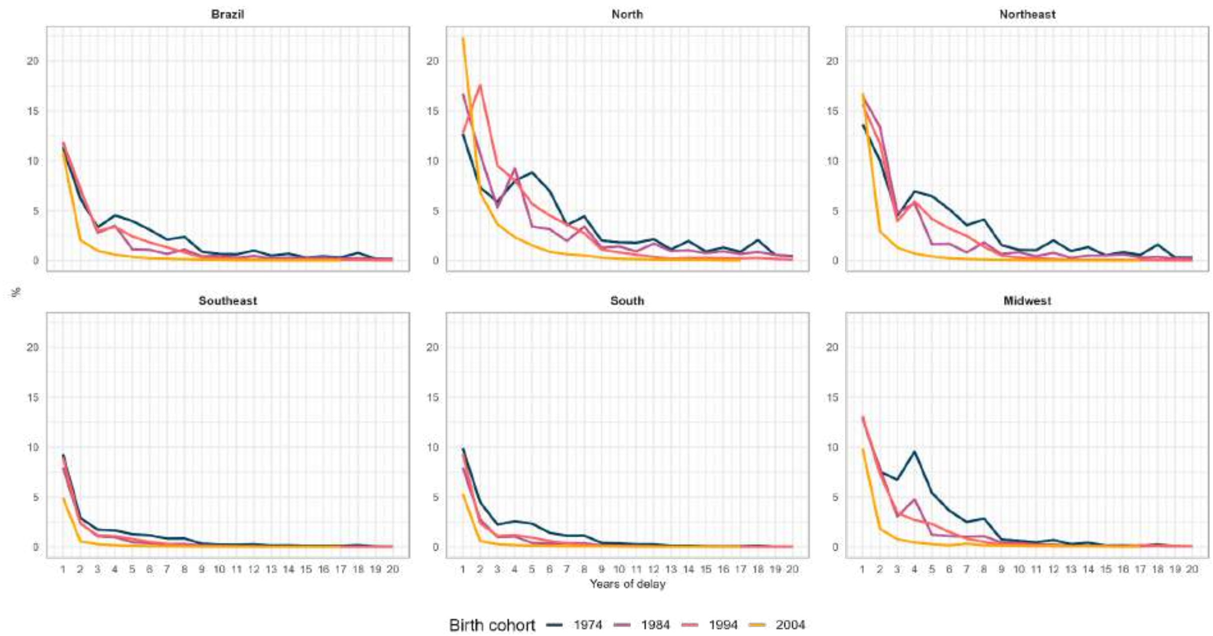
Analyzing the changes in timely and delayed birth registrations over time provides valuable insights. However, to gain a broader perspective, it is important to examine the distribution of registrations based on the years of delay. Figures 5 and 6 present this information separately for each sex and different regions of Brazil. In these plots, the proportion of timely registrations (0 years of delay) is excluded to provide a more precise visualization of the patterns of delayed registrations.

Figure 5: Proportion of Late Birth Registrations by Years of Delay and Birth Cohort (Relative to Total Registrations Within 20 Years of Birth), 1974–2004. Brazil and Brazilian Regions– **Females**



Source: IBGE.

Figure 6: Proportion of Late Birth Registrations by Years of Delay and Birth Cohort (Relative to Total Registrations Within 20 Years of Birth), 1974–2004. Brazil and Brazilian Regions – **Males**



Source: IBGE.

In the type of graph used in Figures 5 and 6, a larger area under the curve, indicating how far the curve is from the graph's axes, is correlated with a higher proportion of delayed registrations. If 100% of births were registered in the year of occurrence, the curve would lie flat on the 0% axis for all years of delay. Across the graphs in these figures, a consistent pattern emerges for all regions: younger cohorts show lower levels of delay, with curves gradually taking on an "L" shape. This pattern suggests reduced variability in registration delays, as most births are registered within a year of the birth year.

While a consistent pattern of smoother, lower, and slightly "L"-shaped curves was observed across all regions over time, notable regional disparities remain. Specifically, the North and Northeast regions show more erratic patterns, particularly among older cohorts, with a higher proportion of delayed registration for longer delays. In the North region, 31.8% of male births and 32.9% of female births from the 1974 cohort were registered six or more years late. In the Northeast, the proportion for both sexes was around 25%. However, substantial improvements were noted over time: by the 2004 cohort, the percentage of births registered six or more years late had sharply declined to about 3% in the North and 0.7% in the Northeast. In contrast, the differences when compared to the Southeast and South regions are significant.

In summary, Brazil has made significant progress in improving timely birth registration. There has been a considerable decline in delayed registrations across all analyzed cohorts and in all states and regions. At the national level, the percentage of births registered in the year they occurred rose from 56.7% for males and 55.7% for females in the 1974 cohort to 84.4% for both sexes in the 2004 cohort. However, this progress has not been consistent across regions. Among the 2004 cohort, the 10 states with the highest prevalence of timely birth registration for both sexes were:

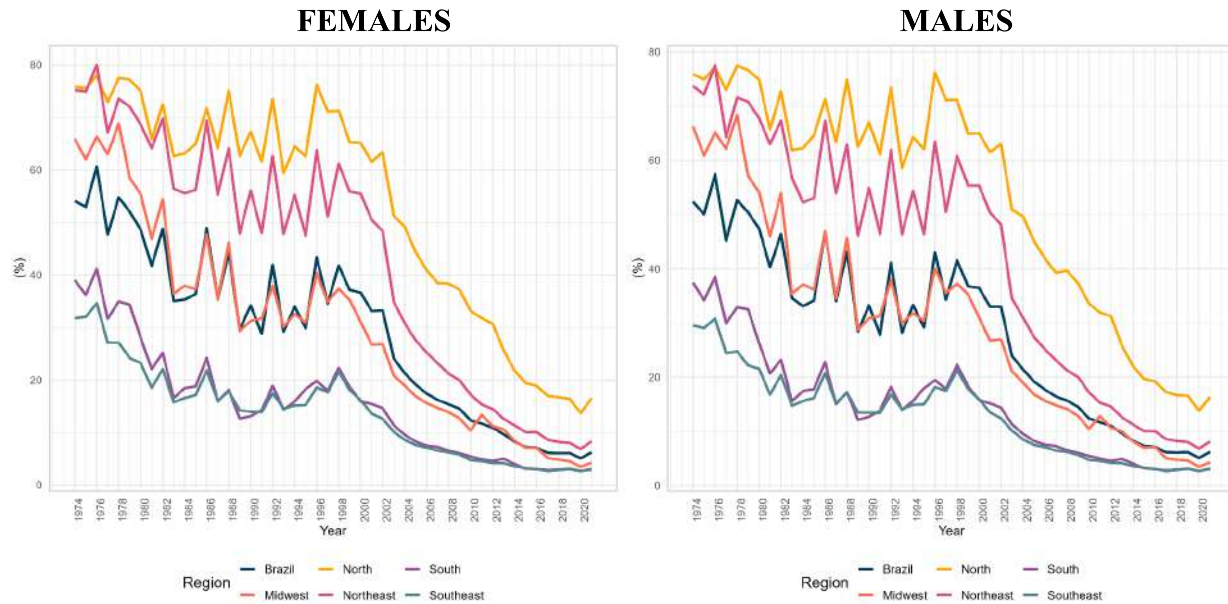
- São Paulo (95.3% for females; 95.3% for males)
- Santa Catarina (94.5% for females; 94.5% for males)
- Paraná (93.5% for females; 93.6% for males)
- Minas Gerais (91.9% for females; 92% for males)
- Distrito Federal (91.8% for females; 92.3% for males)
- Rio Grande do Sul (91.6% for females; 91.5% for males)
- Rio de Janeiro (91.3% for females; 91.4% for males)
- Espírito Santo (91.3% for females; 91.5% for males)
- Goiás (86.9% for females; 87% for males)
- Paraíba (84.1% for females; 84.7% for males)

Notably, the top eight states are all located in the Southeastern and Southern regions, while Distrito Federal, Goiás and Paraíba represent the only states outside those regions. Moreover, there is a substantial gap between the 8th and 10th positions; for instance, the difference in timely registration for females between Espírito Santo (8th) and Paraíba (10th) is 7.2 percentage points.

Period perspective

From a period perspective, the proportion of delayed registrations among total yearly registrations declined sharply. In 1974, 54.1% of female and 52.4% of male births registered that year were from previous years. By 2021, this had declined substantially across all regions, though disparities remained, with the North and Northeast still showing higher rates than the Southeast and South (Figure 7).

Figure 7: Proportion of Delayed Birth Registrations over the years (Relative to Total Registrations in the year) – 1974–2021. Brazil and Brazilian Regions.



Source: IBGE.

The Relationship Between Birth Registration Delays and Old-Age Indicators in Census Data

Table 3 presents Pearson correlation coefficients between indicators of delayed birth registration for the year 1974 and selected old-age demographic indicators from the 2022 census. These correlations are categorized by sex and calculated at the state level.

Pearson correlation coefficients are used to measure both the strength and direction of the relationship between two quantitative variables. The coefficient ranges from –1 to +1. Negative values indicate an inverse relationship, and positive values signify a direct relationship. A coefficient closer to –1 or +1 indicates a stronger association, whereas a value closer to 0 suggests a weak or non-existent association (0 indicates no relationship at all) (Moore et al., 2013).

Table 3 shows a strong positive correlation between delayed birth registrations in 1974 and old-age indicators in 2022, including age ratios (100+/80+ and 100+/50+) and centenarians per 10,000 inhabitants. Correlation coefficients range from 0.48 to 0.81, depending on the variable and sex. In short, states with more delayed registrations in 1974 also tend to report more centenarians relative to older age groups and the total population.

Table 3: Pearson Coefficients Between Delayed Birth Registrations in 1974 and Old-age Indicators by State for Men and Women

Indices	***Women***			***Men***		
	Age ratio 100+/80+ (2022)	Age ratio 100+/50+ (2022)	Centenarians per 10,000 (2022)	Age ratio 100+/80+ (2022)	Age ratio 100+/50+ (2022)	Centenarians per 10,000 (2022)
Proportion of Total Delayed Birth Registrations in 1974	0.64	0.65	0.48	0.78	0.81	0.75
Proportion of Birth Registrations Delayed Over 50 Years (1974)	0.81	0.83	0.83	0.88	0.86	0.86
Proportion of Birth Registrations Delayed Over 30 Years (1974)	0.74	0.75	0.75	0.86	0.85	0.85
Proportion of Birth Registrations Delayed Over 25 Years (1974)	0.75	0.76	0.76	0.86	0.86	0.86

* All values are statistically significant (p-value < 0.05).

* Delayed birth registrations refer to cases where registration occurred in a year later than the actual year of birth.

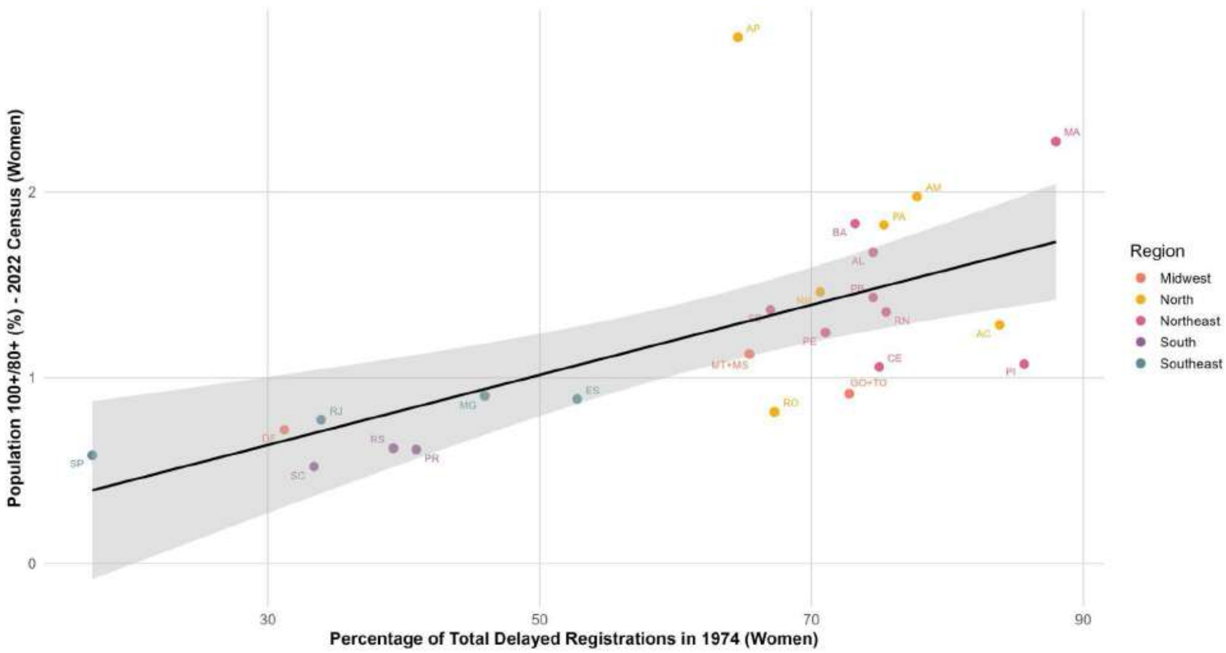
* For example, the "Proportion of birth registrations delayed over 50 years in 1974" in our dataset represents registrations in 1974 for individuals who reported their birth years as being between 1900 and 1924

Although those born in 1974 were not yet centenarians by 2022, 1974 is the earliest year with available registration data. To better align with the centenarian population, we also examined the share of births registered with delays over 50 years in 1974 (that is, individuals who reported being born between 1900 and 1924). Assuming a maximum lifespan of 120 years, the centenarian population in 2022 includes those born between 1902 and 1922.

As Table 3 shows, correlations are even stronger when focusing on extreme delays (50+ years), indicating that states with more very late registrations also report inflated old-age indicators. For example, among men, the correlation between 50+ year delays in 1974 and the 100+/80+ ratio in 2022 reaches 0.88. Scatterplots below illustrate these relationships.

Figures 8 and 9 show the relationship between the 100+/80+ age ratio in 2022 and two indicators of delayed birth registrations for women: the overall proportion of delayed registrations in 1974 (Figure 8) and the proportion of extreme delays (50+ years) (Figure 9). Figures 10 and 11 present the same relationships for men. All graphs include a black OLS regression line.

Figure 8: Relationship Between Age Ratio 100+/80+ (2022 Census) and Percentage of Total Delayed Birth Registrations (1974) by Brazilian States **– Women**

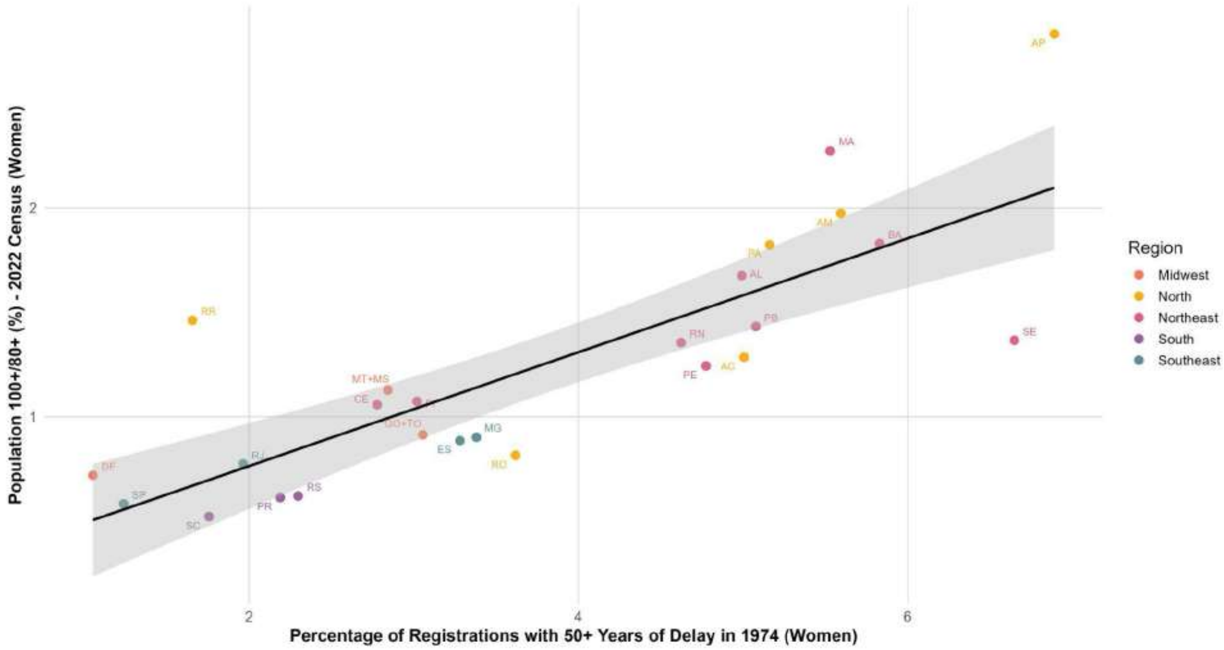


Source: IBGE. Elaborated by the authors.

Among women, Figures 8 and 9 reveal clear regional patterns. In Figure 8, North and Northeast states (yellow and pink) show both higher delayed registrations in 1974 and higher 100+/80+ ratios in 2022. In contrast, states in the South and Southeast, along with the Federal District, cluster in the lower-left corner, reflecting lower delayed registration rates and lower proportions of centenarians. A similar pattern appears in Figure 9 for extreme delays (50+ years), with northern and northeastern states concentrated in the upper-right area and southern states, São Paulo, Rio de Janeiro, and the Federal District in the lower-left.

In both graphs, Amapá and Maranhão stand out as outliers with exceptionally high delayed registration rates and 100+/80+ values. Roraima also deviates, with low extreme delays but a high centenarian ratio. As Giraldelli and Wong (1984) note, in areas with weak registration systems, fewer extreme delays may reflect under-registration rather than improved registration practices.

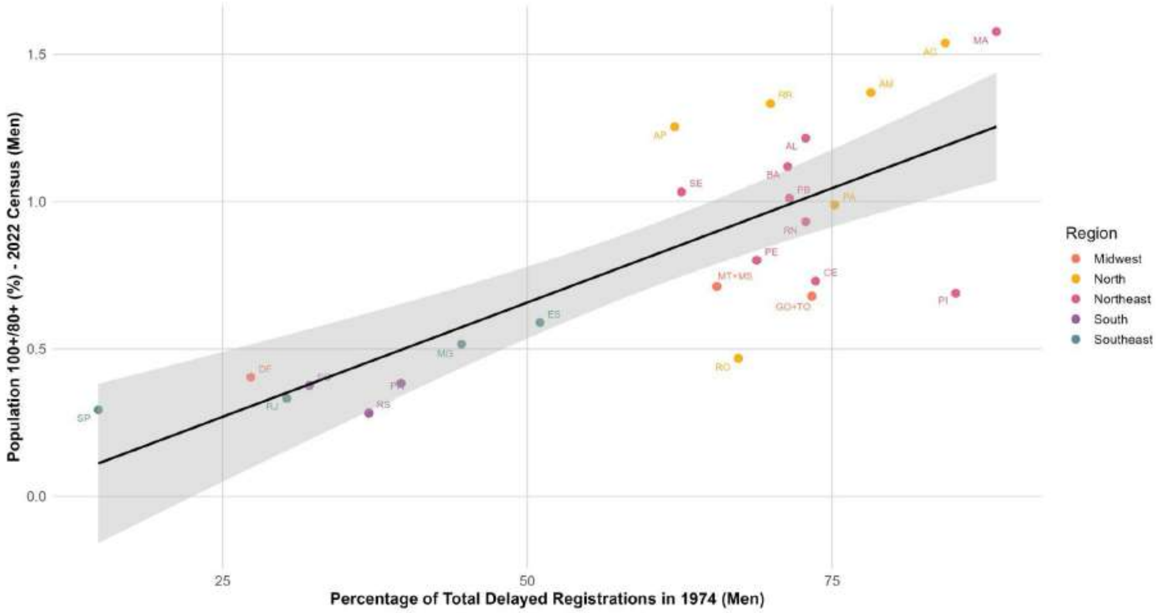
Figure 9: Relationship Between Age ratio 100+/80+ (2022 Census) and Percentage of Birth Registrations with 50+ Years of Delay in 1974 by Brazilian States **– Women**



Source: IBGE. Elaborated by the authors.

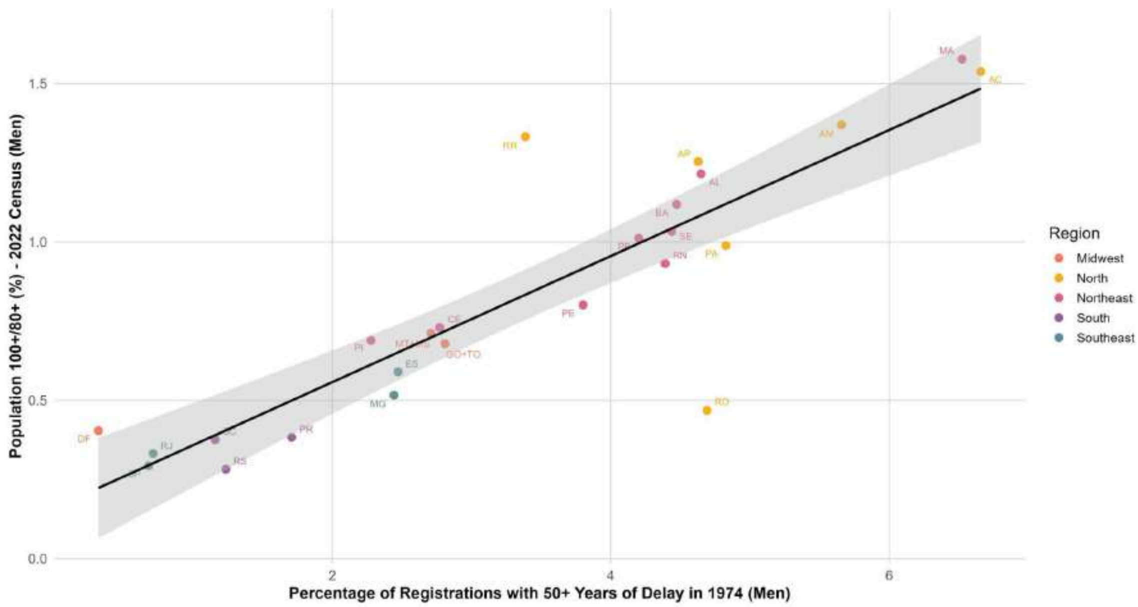
Figures 10 and 11 show that, among men, the relationship between delayed birth registrations in 1974 and the 100+/80+ age ratio in 2022 is more consistent across states, with fewer outliers, especially in Figure 11, which focuses on extreme delays (50+ years). Still, the overall pattern mirrors that observed for women. North and Northeast states cluster in the upper-right areas of both graphs, reflecting high levels of delayed registrations and elevated 100+/80+ ratios. Maranhão stands out again, recording both the highest 100+/80+ ratio and the highest share of total delayed registrations in 1974, along with the second-highest proportion of registrations delayed by over 50 years.

Figure 10: Relationship Between Age Ratio 100+/80+ (2022 Census) and Percentage of Total Delayed Birth Registrations (1974) by Brazilian States **– Men**



Source: IBGE. Elaborated by the authors.

Figure 11: Relationship Between Age ratio 100+/80+ (2022 Census) and Percentage of Birth Registrations with 50+ Years of Delay in 1974 by Brazilian States **– Men**



Source: IBGE. Elaborated by the authors.

These findings underscore the potential long-term implications of historical deficiencies in birth registration systems. It is plausible that states with more severe birth registration problems in the past would continue to show distortions in their old-age population indicators. The data confirm this: states with the most implausibly high old-age indicators are often those that recorded the highest prevalence of delayed registrations in the 1970s, likely reflecting even higher rates in earlier decades. While correlation does not imply causation, these associations remain noteworthy.

The longer the delay between birth and registration, the less reliable the recorded birth information becomes (Hedrich, 1942; AbouZahr et al., 2015; Harbers, 2020). Individuals may unknowingly report incorrect birth dates or deliberately falsify them to access old-age benefits or avoid military obligations, for example (Ewbank, 1981).

5. Discussion

This study examined age data quality in Brazilian censuses (1970–2022), the patterns of delayed birth registration in Brazil, and the relationship between the two. First, we examined the quality of age reporting in Brazilian censuses from 1970 to 2022 at the state level, with a particular focus on the "older adults" population (individuals aged 80 and above). Various indicators were applied, including Whipple's and Myers' indices and population proportions. Additionally, we calculated intercensal survival ratios, analyzed the spatial distribution of the older adults across states, and assessed variations by state, sex, and over time. Building on the hypothesis that age misreporting is associated with weaknesses in Brazil's birth registration system (Turra et al., 2023), we explored patterns of delayed birth registration and its relationship with the data quality indicators first calculated.

Age reporting quality has improved markedly in terms of digit preference, as evidenced by both Whipple and Myers indices. Brazil's Whipple index declined from 126.3 in 1970 (classified as "rough") to below 105 after 1991, indicating "very accurate data" (Grupo de Foz, 2021). Similarly, the Myers index declined by approximately 75% over the same period. However, national averages obscure substantial regional disparities. The North, Northeast, and Midwest regions have consistently lagged behind the South and Southeast, though the gap has narrowed over time.

While these indices provide valuable insights, their limitations must be acknowledged. The Whipple index, calculated for ages 23 to 62 as recommended by the United Nations (1955), captures preference only for ages ending in "0" and "5," providing a limited perspective on overall data quality. In contrast, the Myers index accounts for preferences across all terminal digits and, in this study, was calculated for ages 80 to 99, given the focus on the older adults. Although Myers index results reveal a steady decline across all regions, it is important to note that low Myers index values do not necessarily indicate the absence of age misreporting, as both indices primarily measure digit preference and cannot detect systematic age exaggeration.

To assess systematic age exaggeration, we analyzed distortions in age ratios at older ages, such as the 100+/80+ and 90+/80+ ratios. The findings indicate substantial age exaggeration in Brazil, with a disproportionately high number of centenarians relative to the population aged 80 and above, compared to countries with reliable data, such as Sweden, France, and Italy. For instance, in 2022, the northern state of Amapá recorded a 100+/80+ ratio of 2.15 for both sexes, whereas Sweden's ratio for the same year was only 0.46.

These findings align with Turra et al. (2023), who analyzed Myers index values and 100+/80+ ratios for Brazilian regions using census data from 1980 to 2010 and elderly population counts from the COVID-19 vaccination database (2021). Our results for the 2022 census show higher Myers index values across all regions compared to those observed by Turra et al. (2023) using vaccination records. This suggests that the improved data quality in vaccination records cannot be explained solely by historical trends, as those authors proposed. Instead, the comparison with the most recent census results indicates that requiring proof of age during vaccination campaigns effectively reduced age heaping beyond what would be expected from historical improvements alone. Conversely, our 100+/80+ ratio results for the 2022 census are lower than those from the vaccination database, indicating that relatively fewer centenarians were reported in the census than in the vaccination records. The reasons for this discrepancy remain unclear, especially considering that higher age exaggeration is typically expected in census data rather than administrative records.

In terms of geographic distribution, the prevalence of individuals aged 80 and older and 100 and older increased significantly between 1970 and 2022, with notable spatial disparities. In 1970, the four states with the highest concentrations of men aged 80+ per 10,000 inhabitants were all located in the Northeast. By 2022, the Northeast continued to report the highest concentration of centenarians per 10,000 inhabitants, for both sexes. For instance, in the 2022 census, only three states – Paraíba (4.49), Maranhão (4.66), and Bahia (5.21) – reported four or more centenarian women per 10,000 inhabitants, all in the Northeast. This is particularly striking given the region's comparatively worse socioeconomic conditions. Meanwhile, Santa Catarina, a southern state historically associated with lower mortality rates (IBGE, 2024), reported only 1.21 centenarian women per 10,000 inhabitants (IBGE, 2022).

Beyond relative rates, the absolute number of centenarians in these three states is also remarkable. In 2022, Paraíba recorded 1,330 centenarians (407 men and 923 women), Maranhão reported 2,471 (865 men and 1,606 women), and Bahia reported the highest count, with 5,336 (1,553 men and 3,803 women). Collectively, these states accounted for 9,137 centenarians, representing 24% of Brazil's centenarian population in the 2022 census, despite comprising only 12% of the national population. Notably, Bahia not only had the highest centenarian rate per 10,000 inhabitants but also the highest absolute number, surpassing São Paulo's 5,095 centenarians, despite São Paulo's population being 3.1 times larger (IBGE, 2022). At the national level, discrepancies also persist: while the 2022 census recorded 37,814 centenarians (IBGE, 2022), United Nations estimates for the same year reported only 7,724 (WPP, 2024). These inconsistencies strongly suggest significant age misreporting, particularly in specific regions, which may also affect other age groups.

This uncertainty regarding the true size of the oldest-old population stems not only from the widespread tendency to exaggerate age (Rosenwaike and Preston, 1984; Dechter and Preston, 1991; Preston, Rosenwaike, and Hill, 1996) but is also linked to the historical weaknesses in Brazil's birth registration system. Turra et al. (2023) emphasize that even when individuals report birth dates that match official documents, those dates may be inaccurate due to the widespread lack of reliable birth registration in Brazil's recent past, characterized by high rates of omission and delayed registration (Giraldelli and Wong, 1984).

In this context, we examined patterns and progress in birth registration across Brazilian states, along with associations between registration quality and the age-reporting indicators from the first section. The results reveal significant improvements in Brazil's birth registration system between 1974 and 2021, with a substantial reduction in late registrations across all states, albeit with considerable regional variation. We interpret the increasing prevalence of timely registrations as an indicator of improved system quality, as late registrations are more susceptible to inaccuracies in reported birth dates and places of birth (Hedrich, 1942; AbouZahr et al., 2015; Harbers, 2020). Moreover, we observed a strong positive correlation between delayed birth registration and indicators such as the 100+/80+ age ratio. However, a key limitation is the inability to establish causation. While the correlation is clear, states with implausibly high 100+/80+ ratios also exhibit a higher prevalence of delayed birth registration; the causal pathway remains uncertain. Nonetheless, these findings represent a significant contribution by utilizing novel data to provide a comprehensive cohort and period analysis.

This study makes several novel contributions. First, we provide comprehensive state-level evidence of age misreporting across six censuses (1970–2022), including the recently released 2022 census, which remains underexplored. Second, we analyze previously unpublished state-level birth registration data from IBGE's Research Committee, extending beyond prior work limited to São Paulo (Giraldelli and Wong, 1984). Third, and most importantly, this is the first study to systematically link patterns of delayed birth registration with census age data quality in Brazil, demonstrating the potential long-term implications of historical deficiencies in registration systems.

These findings have significant implications beyond Brazil. Many low- and middle-income countries have experienced similar trajectories of uneven development and civil registration challenges, suggesting that observed demographic patterns, particularly regarding older populations, may reflect administrative data artifacts rather than reality. Researchers and policymakers must exercise caution when interpreting age-based demographic indicators derived from such contexts. The methodology developed in this study offers a replicable approach for assessing the relationship between birth registration quality and census age data accuracy in such settings.

References

AbouZahr, C., De Savigny, D., Mikkelsen, L., Setel, P. W., Lozano, R., Nichols, E., ... & Lopez, A. D. (2015). Civil registration and vital statistics: progress in the data revolution for counting and accountability. The Lancet, 386(10001), 1373-1385.

Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W., & Wachter, K. W. (2018). The plateau of human mortality: Demography of longevity pioneers. Science, 360(6396), 1459-1461.

Bhat, P. M. (1990). Estimating transition probabilities of age misstatement. Demography, 27, 149-163.

Bhatia, A., Krieger, N., Beckfield, J., Barros, A. J., & Victora, C. (2019). Are inequities decreasing? Birth registration for children under five in low-income and middle-income countries, 1999–2016. BMJ global health, 4(6), e001926.

Brill, R. J. (2013). Noncontributory pensions, cash transfers, and documentation in Brazil and Latin America.

Brito, S., Corbacho, A., & Rivas, R. O. (2013). Birth registration: the key to social inclusion in Latin America and the Caribbean.

Campos, N. O. B., & Nascimento Rodrigues, R. (2004). Ritmo de declínio nas taxas de mortalidade dos idosos nos estados do Sudeste, 1980-2000. Revista Brasileira de Estudos de População, 21(2), 323-342.

Carvalho, J. A. M. D., & Garcia, R. A. (2003). O envelhecimento da população brasileira: um enfoque demográfico. Cadernos de Saúde Pública, 19, 725-733.

Coale, A. J. (1964). How a population ages or grows younger. in Ronald Freedman (ed.),. Population: the Vital Revolution (Garden City, New York, 1964), 47-58.

Coale, A. J., & Kisker, E. E. (1986). Mortality crossovers: Reality or bad data?. Population studies, 40(3), 389-401.

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.

Dunning, C., Gelb, A., & Raghavan, S. (2014). Birth registration, legal identity, and the post-2015 agenda. Washington, DC: Center for Global Development.

Ewbank, D. C. (1981). Age misreporting and age-selective underenumeration: Sources, patterns, and consequences for demographic analysis (Vol. 4). National Academy Press.

Gavrilov, L. A., & Gavrilova, N. S. (2019). Late-life mortality is underestimated because of data errors. PLoS biology, 17(2), e3000148.

Giraldelli, B. W., & Wong, L. R. (1984). O comportamento do registro atrasado de nascimentos (RAN) no Estado de São Paulo: uma tentativa de correção do sub-registro. Anais do IV Encontro Nacional de Estudos Populacionais.

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

Grupo de Foz (Ed.). (2021). Métodos demográficos: uma visão desde os países de língua portuguesa. Blucher Open Access.

Grushka, C. O. (1996). Adult and Old Age Mortality in Latin America: evaluation, adjustments and a debate over a distinct pattern. University of Pennsylvania.

Harbers, I. (2020). Legal identity for all? Gender inequality in the timing of birth registration in Mexico. World Development, 128, 104778.

Hedrich, A. W. (1942). Delayed Birth Registration. American Journal of Public Health and the Nations Health, 32(4), 365-368.

HMD. Human Mortality Database. (2025). Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France). Available at www.mortality.org.

Hunter, W., & Brill, R. (2016). "Documents, please": advances in social protection and birth certification in the developing world. World Politics, 68(2), 191-228.

Hunter, W. (2018). Identity documents, welfare enhancement, and group empowerment in the Global South. The Journal of Development Studies, 55(3), 366-383.

IBGE. (2000). Demographic Census. Rio de Janeiro: IBGE.

IBGE. (2002). Estatísticas do Registro Civil 2000. Vol 27. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística; 2002.

IBGE. (2010). Demographic Census. Rio de Janeiro: IBGE.

IBGE. (2011). Estatísticas do Registro Civil 2010. Vol 37. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística; 2011.

IBGE. (2021). Data on Registrations by Year of Birth vs Year of Registration. Provided by the "Diretoria de Pesquisas".

IBGE. (2022). Demographic Census. Rio de Janeiro: IBGE.

IBGE (2024). Projeções da População do Brasil e Unidades da Federação: 2000-2070. IBGE, Diretoria de Pesquisas. Rio de Janeiro: IBGE.

IBGE. (2024b). Estatísticas do Registro Civil 2022. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística; 2024.

Jayaraman, J., Roberts, G. J., Wong, H. M., McDonald, F., & King, N. M. (2016). Ages of legal importance: implications in relation to birth registration and age assessment practices. Medicine, Science and the Law, 56(1), 77-82.

Lundquist, J. H., Anderton, D. L., & Yaukey, D. (2014). Demography: the study of human population. Waveland Press.

Minnesota Population Center (2024). Integrated Public Use Microdata Series, International: Version 7.5. Minneapolis, MN: IPUMS. doi:10.18128/D020.V7.5.

Moore, D. S., Notz, W. I, & Flinger, M. A. (2013). The basic practice of statistics (6th ed.). New York, NY: W. H. Freeman and Company.

Myers, R. J. (1966). Validity of centenarian data in the 1960 census. Demography, 3(2), 470-476.

Myers, R. J. (1976). An instance of reverse heaping of ages. Demography, 13(4), 577-580.

Nepomuceno, M.R. and Turra, C.M. (2020). The population of centenarians in Brazil: Historical estimates from 1900 to 2000. Population and Development Review 46(4): 813–833. doi:10.1111/padr.12355.

Newman, S. J. (2024). The global pattern of centenarians highlights deep problems in demography. medRxiv, 2024-09.

Oeppen, J., & Vaupel, J. W. (2002). Broken limits to life expectancy. Science, 296(5570), 1029-1031.

Olshansky, S. J., Carnes, B. A., & Cassel, C. (1990). In search of Methuselah: estimating the upper limits to human longevity. Science, 250(4981), 634-640.

Ortega, A., & García, V. (1985). Estudio experimental sobre la mortalidad y algunas características socioeconómicas de las personas de la tercera edad: informe de la investigación efectuada en los cantones de Puriscal y Coronado, 3-20 junio 1985.

Palloni, A., Beltrán-Sánchez, H., & Pinto, G. (2021). Estimation of older-adult mortality from information distorted by systematic age misreporting. Population studies, 75(3), 403-420.

Preston, S. H., Elo, I. T., Rosenwaike, I., & Hill, M. (1996). African-American mortality at older ages: Results of a matching study. Demography, 33(2), 193-209.

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. Blackwell Publishers. Oxford, UK.

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.

Rau, R., Ebeling, M., Peters, F., Bohk-Ewald, C., & Missov, T. I. (2017, January). Where is the level of the mortality plateau. In Living to 100 Symposium of the Society of Actuaries, Orlando, FL.

Reis, C. S., & Turra, C. M. (2016). Distribuição espacial dos centenários no Brasil: uma análise exploratória da qualidade dos dados dos censos de 2000 e 2010. Revista Espinhaço.

Richman, R.D. (2017). Old age mortality in South Africa, 1985–2011. Master's thesis. University of Cape Town.

Rosenwaike, I., & Preston, S. H. (1984). Age overstatement and Puerto Rican longevity. Human Biology, 503-525.

Siegel, J.S. and Swanson, D.A. (Ed.). (2004). The methods and materials of demography. Second edition. New York: Elsevier/Academic Press.

Turra, C. M. (2012). Os limites do corpo: a longevidade em uma perspectiva demográfica. Revista da Universidade Federal de Minas Gerais, 19(1 e 2), 156-181.

Turra, C. M., Fernandes, F., Calazans, J. A., & Nepomuceno, M. R. (2023). Age reporting for the oldest old in the Brazilian COVID-19 vaccination database. Demographic Research, 48, 829-848.

UNICEF. (2005). The 'rights' start to life: a statistical analysis of birth registration. Unicef.

UNICEF. (2013). EVERY CHILD'S BIRTH RIGHT: Inequities and trends in birth registration. Unicef.

United Nations (1955). Manual II Methods of Appraisal of Quality of Basic Data for Population Estimates. New York: United Nations.

Wong, L. R., & Turra, C. M. 2007. O Sub-Registro de Nascimentos No Brasil. Inter-American development bank project ATN/NS-8915-RS).

World Population Prospects (2024). The 2024 Revision. United Nations Department of Economic and Social Affairs 2024. <https://population.un.org/wpp>