Age reporting in the Brazilian COVID-19 vaccination database: What can we learn from it?

Extended Abstract

Cássio M Turra (Cedeplar, UFMG), Fernando Fernandes (Cedeplar, UFMG), Julia Calazans (Cedeplar, UFMG), Marilia Nepomuceno (Max Planck Institute for Demographic Research)

Introduction

Age is a crucial variable for social and health sciences. However, academic research and public planning often neglect the quality of age reporting. The consequences of not measuring age correctly are manifold, including errors in mortality rates and population estimates and projections, particularly at older ages (Preston et al. 1999). Age misreporting also impacts vaccination policy. During the COVID-19 pandemic, this topic became even more relevant since age was one of the main variables discussed in vaccine prioritization strategies (Castro and Singer 2021, Goldstein et al. 2021).

Age misreporting varies across data sources, countries, population subgroups, and periods (Coale and Kisker 1986, Elo and Preston 1994, Dechter and Preston 1991, Nepomuceno and Turra 2020, Gomes and Turra 2009). Research has associated this data issue with different factors, including delayed birth registration, which precludes individuals from knowing their actual age, response errors from memory recall problems, proxy reporting, and low numeracy and literacy. Many household surveys and administrative records now include a question about the birth date instead of age in completed years, which has improved data quality, mitigating age heaping. Still, in places where there are insufficient quality birth registration data, recalling the actual birth date may remain a challenge, particularly among the elderly of low SES.

Brazilian data are not free of age misreporting. Turra (2012) has claimed that the pattern of mortality at older ages in Brazil, characterized by rates relatively lower than in high-income countries, reflects age exaggeration. Recent findings have confirmed his argument (Nepomuceno & Turra, 2020; Di Lego, Turra, e Cesar, 2017). Also, the excess of centenarians in official Brazilian data is consistent with age misreporting due to delayed birth registration during the 19th and twentieth centuries (Nepomuceno & Turra 2020; Gomes & Turra 2014). Quality of age reporting is probably not limited to the oldest old and may affect population estimates and deaths at younger ages. Studies on adult mortality in Latin America suggest this is perhaps Brazil's case (Dechter and Preston 1991; Rosenwaike and Preston 1984; Grushka 1996). However, we still do not know at what age vital statistics and census data become seriously flawed in Brazil.

As vital registration improves for the new birth cohorts, one can expect age reporting to become intrinsically more accurate. Still, it will take some time to get precise age distributions at advanced ages. In the absence of high-quality vital records, historical census data and parochial records have helped demographers identify the correct age of adults by tracking events and reports from the first years of life. Unfortunately, historical census microdata are missing for Brazil, and parochial records are selective and not fully organized in a national database. The remaining option is to use statistical and demographic methods to adjust the current population and death counts by age (Palloni et al. 2016, Carrier and Farrag, 1959, Siegel and David 2004.). But estimation also involves uncertainty since the actual patterns are unknown, and data errors may confound with other factors such as mortality selectivity.

Therefore, any new database that keeps age or birth date records can be precious for population estimates in the Brazilian context. Systematic comparisons across old and new datasets offer fresh clues about any differences with the official figures. For example, data collected from high SES subgroups may reveal similarities in the estimated age patterns compared to those calculated from other countries known for their high-quality data (Di Lego, Turra, e Cesar, 2017). On the other hand, these comparisons can provide a critical assessment of the quality of the new datasets. It is a two-way street in terms of reducing our ignorance about age and other demographic distributions.

During a global public health crisis, the demand for high-quality information increases, and there are opportunities and challenges for building new datasets. The Covid-19 pandemic has resulted in intense activity on data collection, cleaning, and quality monitoring. After the Sars-CoV-2 hit Brazil,

gathering the number of cases, hospitalizations, and deaths has become vital for curbing the virus. Despite the many difficulties faced by local authorities, the collaborative effort of public institutions, scholars, and civil society enabled compiling information from municipalities, helping to keep track of the pandemic. The start of the vaccination campaign on January 18th of 2021, marked the collection of additional national data. The National Plan for Operationalization of the Vaccine against Covid-19, implemented by the Federal Government, has defined individuals 60 and older as one of the priority groups (REF). The reason is the higher mortality among the elderly (Richardson et al. 2020). Therefore, age is a crucial variable for vaccination, and every vaccinated person needs to be registered. In addition to the elderly, the priority group includes individuals ages 18 to 59 living with specific comorbidities, healthcare workers, traditional peoples and communities (quilombolas, indigenous people, and riverside dwellers), the deprived of liberty, education workers, and other workers considered essential. Following these prioritization definitions, the Brazilian government used different data sources to estimate the target populations. According to official documents, the whole priority group amounts to 77,279,644 people, from which 30,197,052 are individuals ages 60 and older.

The national vaccination plan requires individuals to show identity cards to prove their age when getting vaccinated. That is not the case of most household surveys, including census data and other sources used to estimate the target populations in 2021. Proof-of-age documentation cannot solve wrong ages from delayed birth registration. Still, it can mitigate age misreporting when response errors originate from memory recall problems, proxy reporting, or low numeracy and literacy. Therefore, whereas some individuals may avoid getting immunized, we expect more accurate age distributions in the vaccination registration system than in the official population estimates. We hypothesize that older population groups are overestimated in the National vaccination plan due to the known age exaggeration at these age groups in Brazil (Turra 2012, Nepomuceno and Turra 2020). If this is the case, vaccination coverage rates by age may be misleading and should be taken with a grain of salt.

Missing the population target is unfortunate, particularly in a country suffering from high death rates from Covid-19 and lacking vaccine doses to distribute for other age groups. In this article, we try to cast some light on these issues by looking at age distributions of men and women 80 years and older of age registered in the vaccination system. Limiting the age range is necessary since the vaccination program is still ongoing. The Brazilian campaign started with the oldest groups making the data more stable for ages older than 80. We calculate different demographic indicators using data from the vaccination registration system and compare them to those from the target population estimates (National Vaccination Plan). We also compare them with indicators calculated from the 2010 Brazilian census, population projections prepared by both the Brazilian Bureau of Census (IBGE) and the United Nations (U.N) for 2021, and countries with high-quality population data. Further, we look at results by Brazilian states since age reporting is associated with education and other socio-economic conditions, and results may vary across the territory affecting the regional distribution of vaccines. Our analyses make two critical contributions to the demographic and public health literature. First, it offers new critically assessed estimates of the Brazilian elderly population based on the vaccination registration system. Second, it uncovers potential problems in measuring vaccination coverage rates with the current target population estimates.

Materials

Data

We drew data from the Brazilian Ministry of Health open microdata registers for the national Covid-19 vaccination campaign (Brasil 2021a). We used vaccination registers recorded from the beginning of the campaign (January 2021) until May 06th, 2021 (the date we download the data). The database comprises anonymized information on both first and second doses of any Covid-19 vaccine. It also includes data on anonymized document I.D., anonymized patient I.D., dose (first or second), age, date of birth, sex, race, vaccination group (e.g., health worker, age group), city and state of vaccination and residence, vaccine name and developer. Our study considered only the population vaccinated with the first dose of any Covid-19 vaccine until May 06th, 2021 for Brazil and its states by sex and ages 80 to 110 and older. The initial data comprised 45,263,837 records, of which 272,466 were either duplicated (265,598), triplicated (6,840), or quadruplicated (28). Therefore, after cleaning 135,086 "deduplicated" records, the data consisted of 45,126,457 records. From this total, we excluded records that are the second dose (14,552,061), had missing dose (242), missing state of residence (248,360), missing sex

(48), vaccination date before January 01st, 2021 (14), and whose age on May 01st, 2021 (calculated from the date of birth) were equal or above 120 years (6,121). For most individuals 120 and older (5,975), the date of birth is December 31st, 1899. We believe this is the system default for unknown birth data, but the information needs further investigation. Although old-age individuals may have died between the beginning of the vaccination campaign and May 06th, 2021, we cannot identify who, among the vaccinated, have died during the first four months. Therefore, we ignored any deaths and estimated the vaccinated population as of May 01st, 2021. Our final data consisted of 30,401,664 individuals 80 years and older, 17,803,151 females and 12,598,513 males.

We compared the vaccinated population with several other estimates for the Brazilian elderly, including the target population defined by the Brazilian Ministry of Health (MoH) national plan for vaccination against Covid-19 (Brasil 2021b), the Brazilian resident population in 2010 (IBGE 2011), and population projections prepared by the Brazilian Institute of Geography and Statistics, IBGE (IBGE 2018) and the United Nations (United Nations 2019). We calculated population distributions by sex and 5-year age groups (up to 100+) to make them comparable across datasets whenever possible. Regrettably, the target population presented by the MoH is for both sexes only and ages up to 90 and older. The IBGE projections are also limited to the 90+ age group but offer results for different Brazilian states. To make the population projections (IBGE and U.N.) comparable to the vaccinated population, we adjusted the figures to May 01st, 2021. We applied exponential interpolation by 5-year age groups and sex, using the projections for July 1st, 2020, and July 1st, 2021. Additionally, we reduced the interpolated population by the proportion of excess deaths during 2020 due to the COVID-19 pandemics, calculated by age, sex and state of residence.

We further assessed the Brazilian population estimates against data from countries known by high-quality population enumerations at old ages. The group includes Japan (2005, 2010, and 2015), Sweden (from 1992 to 2019), and Switzerland (from 2002 to 2019), by 5-year age groups (up to 110+) and sex - raw data from *Human Mortality Database* (2021).

Demographic indicators

In addition to population estimates by age and sex, we calculated age and sex ratios to compare the accuracy of the age distribution of the vaccinated population with estimates from the other datasets.

Preliminary Results

Table 1 compares the population estimates for Brazil by data source, sex, and age groups. In agreement with our initial hypothesis, the vaccinated population is smaller than all population projections for Brazil. Compared to the MoH target estimates, the difference is about 276,000 people or 7% and is concentrated at ages older than 90 (26%). The IBGE projected population is also larger for both sexes and all age groups. There are more than 400,000 individuals ages 80 and older in the IBGE projection compared to the vaccination records. The relative difference varies from about 8% for ages 80-89 to 19% for 90 and older, suggesting age distribution issues, as in MoH estimates. The results are somewhat different for the U.N. population estimates. The total size of the projected population is 5% larger (210,000) than the number of vaccination records. Also, the age distributions are much more comparable. The ratios 90+/80+ are 0.171 and 0.167, respectively, for the vaccinated and the U.N. populations.

The lower number of centenarians estimated by the U.N. than vaccinated, 13 thousand individuals (35%), is one evidence of better-quality projections by the U.N. for the very old. As aforementioned, the vaccination records are not free of age misreporting. Individuals who have wrong dates of birth in their I.D. cards will be recorded with the wrong ages, explaining why the number of vaccinated centenarians is larger. However, the age distribution of the vaccinated population is more like the U.N. one than the IBGE and MoH estimates. This pattern confirms our hypothesis that I.D cards helped solve some reporting errors present in the Brazilian data. Although there are no separate estimates for nonagenarians and centenarians in the IBGE and MoH data, the 90+/80+ ratios suggest a typical pattern of age exaggeration, which would probably result in an even larger number of centenarians in these two datasets than in the vaccination records.

In the final version of the paper, we will explore the regional patterns of population enumeration at older ages and provide additional evidence of higher data quality for the vaccination records

compared to the other Brazilian estimates. The paper will also discuss the implication of our analysis for the demographic projections and the Covid-19 vaccination plan in Brazil.

Table 1 – Population estimates, Brazil, by sex and data source, ages 80 and older

		Population Estimates			Ratios		
Brazil	Vaccine Records as of May 6h, 2021 (1)	Health Ministry Target Estimates (2)	IBGE Population Projections (3)	United Nations Population Division (4)	(2)/(1)	(3)/(1)	(4)/(1)
Both Sexes							
80-89	3,452,562	3,547,173	3,738,288	3,644,967	1.03	1.08	1.06
90-99	675,387			706,363			1.05
100 +	36,855			23,804			0.65
90 +	712,242	893,873	849,229	730,167	1.26	1.19	1.03
80 +	4,164,804	4,441,046	4,587,517	4,375,134	1.07	1.10	1.05
Ratio 90+/80+	0.1710	0.2013	0.1851	0.1669	1.18	1.08	0.98
Women							
80-89	2,106,209		2,270,134	2,234,255		1.08	1.06
90-99	454,925			486,494			1.07
100 +	26,486			17,502			0.66
90 +	481,411		573,811	503,996		1.19	1.05
80 +	2,587,620		2,843,945	2,738,251		1.10	1.06
Ratio 90+/80+	0.1860		0.2018	0.1841		1.08	0.99
Men							
80-89	1,346,353		1,468,154	1,410,712		1.09	1.05
90-99	220,462			219,869			1.00
100 +	10,369			6,302			0.61
90 +	230,831		275,418	226,171		1.19	0.98
80 +	1,577,184		1,743,572	1,636,883		1.11	1.04
Ratio 90+/80+	0.1464		0.1580	0.1382		1.08	0.94

References

Brasil. 2021a. “Campanha Nacional de Vacinação contra Covid-19 - Registros de Vacinação COVID19 - Open Data.” openDataSUS. 2021. https://opendatasus.saude.gov.br/dataset/covid-19-vacinacao/resource/ef3bd0b8-b605-474b-9ae5-c97390c197a8.

———. 2021b. *Plano Nacional de Operacionalização Da Vacinação Contra a Covid-19*. 6th ed. Brasília, DF.

Human Mortality Database. 2021. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). www.mortality.org or www.humanmortality.de.

Carrier, N.H. Farrag, A. M. (1959) The reduction of errors in census populations for statistically underdeveloped countries, Population Studies, 12:3, 240-285.

Coale, A.J.; Kisker, E.E.1986. Mortality crossover: reality or bad data? Population Studies, Vol. 40, n.3

Dechter, Aimée R., and Samuel H. Preston. 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, Vanessa, Cássio M. Turra, and Cibeli Cesar. 2017. “Mortality Selection among Adults in Brazil: The Survival Advantage of Air Force Of□cers.” Demographic Research 37(1): 1339–1350

Goldstein, J.R., Cassidy, T., Wachter, K.W. 2021. Vaccinating the oldest against COVID-19 saves both the most lives and most years of life. PNAS 118 (11)

IBGE. 2018. *Projeções Da População: Brasil e Unidades Da Federação - Revisão 2018*. 2nd ed. Séries Relatórios Metodológicos, volume 40. Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais. https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.htm.

Nepomuceno, M.; Turra, C.M. 2020. The population of centenarians in Brazil: Historical etimates from 1900 to 2000.Population and Development Review. 46(4): 813-833.

Palloni, A.; Pinto, G. Beltran-Sanchez. 2016. Estimation of Life Tables in the Latin American Data base (Lambda): adjustments for relative completeness and age misreporting. Manuscript.

Preston, S.H.; Elo, I.T., Stewart, Q. 1999. Effects of Age Misreporting on Mortality Estimates at Older Ages.Population Studies, Vol. 53, No. 2, pp. 165-177

Richardson, S. et al. 2020. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. JAMA 323(20)

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

United Nations. 2019. *World Population Prospects 2019: Online Edition*. New York: United Nations, Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp/.