



MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2024-017 | July 2024
<https://doi.org/10.4054/MPIDR-WP-2024-017>

Forecasting Life Expectancy in São Paulo City, Brazil, amidst the COVID-19 Pandemic

Maria Miranda | miranda@demogr.mpg.de
Cássio Turra
Ugo Filippo Basellini

This working paper has been approved for release by: Emilio Zagheni (sezzagheni@demogr.mpg.de),
Head of the Laboratories of Migration and Mobility and Population Dynamics and Sustainable Well-Being.

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review. Views or opinions expressed
in working papers are attributable to the authors and do not necessarily reflect those of the Institute.

Forecasting Life Expectancy in São Paulo City, Brazil, amidst the COVID-19 Pandemic

Maria Miranda^{1*}, Cássio Turra², Ugofilippo Basellini¹

¹ Max Planck Institute for Demographic Research, Rostock, Germany. ² Center for Development and Regional Planning - CEDEPLAR, UFMG, Belo Horizonte, Brazil.

* Corresponding author: miranda@demogr.mpg.de

June 24, 2024

Abstract

Background:

The COVID-19 pandemic has significantly increased mortality rates, disrupting historical trends and making it challenging to forecast future life expectancy levels. São Paulo, the first city in Brazil to report a COVID-19 case and death, saw a decrease of over four years in life expectancy at birth for males and over three years for females between 2019 and 2021. São Paulo has been at the forefront of the demographic transition in the country and experienced a nonlinear mortality decline over the 20th century. The city's historical mortality trajectory and the disruptive effects of COVID-19 have introduced challenges to mortality forecasting.

Methods:

In this study, we used a unique dataset starting from 1920 to forecast life expectancy in São Paulo until 2050 using the Lee-Carter (LC) and Lee-Miller (LM) methods. Mortality rates were obtained from a combination of deaths gathered by the SEADE Foundation

(SEADE) and population collected by the Brazilian Institute of Geography and Statistics (IBGE). To mitigate the dependency on the fitting period's choice and better incorporate the effects of the recent mortality shock, we used different baseline periods, using all years from 1920 to 1995 as the starting year of the analysis and six scenarios for post-pandemic mortality levels. Additionally, we used a simulation approach for the time-index parameter to calculate prediction intervals. Based on 73,200 simulations for each year between 2023 and 2050, we synthesized the resulting life expectancy forecasts into median values and 95% prediction intervals (PI).

Results:

By 2050, we predict that life expectancy at birth in São Paulo will reach approximately 81.5 years for men and 88.3 years for women. Also, within the 95% PI, we estimated that by 2045, male life expectancy could reach the levels of best-performing countries.

Conclusions:

Our approach is among the first attempts to forecast mortality in the presence of shocks. Additionally, by evaluating different baseline periods, we advocate for the adoption of more accurate forecasting strategies, particularly in contexts of recent mortality decline. These findings provide valuable resources for policymakers and researchers working to address public health challenges arising from the pandemic and plan for the future well-being of many populations.

Keywords: mortality forecasting; covid-19; Lee-Carter; Lee-Miller; São Paulo.

BACKGROUND

São Paulo has been at the forefront of the demographic transition in Brazil. Over the 20th century, the health transition has evolved in the city, improving survival levels and changing mortality patterns [1]. Improvements in public health through the control of diseases, the provision of water, sanitation, and health services, and the combat of violence reduced mortality differently across age groups in the past century, resulting in a non-linear pace of mortality decline [2]. More recently, the effects of COVID-19 in 2020-2022 added more complexity to the historical trajectories of the health transition.

The COVID-19 pandemic triggered an unprecedented rise in mortality globally over the past few years. Such impact was significant enough to change historical trends, resulting in reduced life expectancy and a surprising number of excess deaths for some populations. The discrepancies across countries were mainly due to differences in the age structure [3] and context-specific factors, including socioeconomic conditions, chronic disease prevalence, and public policies to combat the virus spread [4].

Estimates show declines of more than four years in life expectancy at birth between 2019 and 2021 in Bolivia, Botswana, Lebanon, Mexico, Oman, and the Russian Federation [5]. Similar analyses show substantial life expectancy reductions in Eastern Europe and the United States between 2019 and 2021, but not among the Western European countries [6].

In Brazil, the estimated reduction in life expectancy was 1.3 years between 2019 and 2020 [7]. Until May 2021, Brazil was considered the second country with the most notable excess deaths after COVID-19 [8]. From March 15 to June 6, 2020, there were over

62,000 excess deaths caused by diseases observed, representing a 22% increase in overall mortality over 12 weeks [9]. Due to the concentration of the population in major cities, the impact of the pandemic on mortality was more pronounced in urban areas, particularly in the capital cities. São Paulo city, the state's capital, recorded a disproportionately higher number of excess deaths, accounting for 36% (6,208) above the expected average from March 15 to June 6, 2020 [9]. By December 2022, the city had confirmed over 1.1 million cases and more than 44,000 deaths due to COVID-19 [10].

Nevertheless, estimating the pandemic's impact on mortality remains challenging, even when considering measures such as excess mortality, which mitigates the effect of misidentification of COVID-19 as the underlying cause of death. Age-specific case counts depend highly on the testing capacity, testing strategy, and differences in the definition of cases across sources and over time, especially in low- and middle-income countries [11]. Also, due to the behavioral changes induced by the pandemic and the differential fatality rates associated with the presence of comorbidities, establishing a causal relationship between COVID-19 and other causes of death occurring after 2020 becomes complex [12]. It is possible that the rates of these chronic ailments may have or will decrease if patients succumb to COVID-19, a process known as the "harvesting effect" [13; 12]. Also, due to the reduced social and work activities, deaths from specific external causes may have declined during the pandemic [12]. Conversely, deaths from other causes may have increased due to several aspects, including inadequate care in clinics and hospitals because of shortages of equipment, staff, and space, worsening of comorbidities owing to the effects of COVID-19, and other factors [14].

In summary, it is challenging to assess the impact of COVID-19 on mortality due to its interaction with other causes of death [12] and data issues [11]. Additionally, it is

uncertain when mortality levels will return to pre-pandemic levels. Despite the uneven distribution of healthcare infrastructure and ongoing cuts in health spending in Brazil, the public health system has remained operational during the COVID-19 pandemic [15]. In January 2021, the vaccination campaign started in the country, with São Paulo being the first city to vaccinate. The campaign encountered several challenges, including political instability, lack of national planning, delays in securing supplies such as syringes and needles, and vaccine shortages [16; 15]. However, given Brazil's historical culture of immunization, more than 80% of Brazilians aged six months and older completed the primary vaccination schedule as of 2022 [17]. This high acceptance of immunization among the population led to a large reduction of COVID-19 deaths in 2022 (75,429) compared to 2021 (427,629), despite the similar number of identified cases in the two years: 14,768,820 in 2021 and 14,096,655 in 2022 [18]. Considering this rapid decline in COVID-19 deaths, it is plausible that mortality levels will return to pre-pandemic levels in the coming years. However, there remains uncertainty regarding the mortality level's return.

Estimating and forecasting mortality tendencies is essential for designing public policies, assessing health insurance premiums, and examining the solvency of pension systems. It is also a way to indicate levels of well-being across nations, regions, and other population subgroups. However, to estimate future trends, it is necessary to have robust enough historical data to extrapolate mortality levels for the following decades. Additionally, forecasts must consider the uncertainty caused by variations in the observed trend, such as mortality shocks and non-linear declines. São Paulo may be the only place in Brazil where a comprehensive historical series of age-specific mortality rates by sex has been documented, tracing back to 1920. The data quality of this historical series has been

examined before and proved reliable enough to allow demographic exercises, even though it is not error-free [19, 20, 21].

In this study, we estimate future life expectancies in São Paulo city until 2050, incorporating the uncertainty caused by COVID-19 and non-linear changes in mortality trends over the last century. To do this, we use Lee-Carter [22] and Lee-Miller [23] methods with different base periods, starting between 1920 and 1995 until 2022, and six scenarios for post-pandemic mortality levels. Death rates were obtained from the combination of deaths gathered by the SEADE foundation (SEADE) and population collected by the Brazilian Institute of Geography and Statistics (IBGE) data.

METHODS

The Lee-Carter (LC) is a stochastic method developed by Ronald D. Lee and Lawrence R. Carter in 1992 to model and forecast age-specific death rates, combining demographic mortality and time series models [22]. The model was proposed as follows:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

where $\ln(m_{x,t})$ denotes the logarithm of the death rates by age group x at year t ; a_x describes the average shape of the mortality by age; b_x describes the extent to which mortality at age x changes given the overall temporal change in the general level of mortality; k_t is the mortality level index at time t ; and $\varepsilon_{x,t}$ is an error term reflecting residual age-specific influences not captured by the model.

In the original version, Lee and Carter [22] proposed an adjustment on the k_t parameter, so the number of deaths estimated by the method using fixed a_x and b_x would match the ones observed each year t . After this adjustment, they proposed forecasting the k_t s using a random walk process with drift. This time series model can be expressed as follows:

$$k_t = c + k_{t-1} + u_t$$

Where c is the drift and u_t is a white noise process with mean zero and variance σ^2 .

Lee and Carter [22] introduced an innovative and straightforward approach to forecasting mortality rates, yet the method has limitations and disadvantages. The method assumes that the changes in the mortality pattern are constant over time (fixed b_x). This premise fails to account for the possibility that the decline may vary in pace for specific age groups over time. Especially for long-term projections, it is unlikely that such a pattern would hold proportionally to all age groups for more than some decades into the future [24].

The method also relies on \hat{k}_t s linearity and the fitting period's choice. The proposed random walk with drift time series model, as suggested by Lee and Carter [22], is linear and passes through the first and last point of the series [25]. However, mortality levels do not always decrease linearly, particularly among populations considered behind in the demographic transition over the last century, as shown before in the case of São Paulo. Consequently, different choices of the baseline period can lead to forecasts that have various rates of improvement.

Another limitation of the model is the error in the last year of the fitting period (jump-off year), in which the fitted death rates would not match the observed death rates. Given that a_x is considered the mean of all the $\ln(m_{x,t})$ over time, previous years with higher

mortality levels would also influence the predicted levels, leading to higher forecasted mortality levels that don't necessarily reflect the recent trends. Finally, the LC method does not consider information on future trends [26]. This is particularly relevant in the context of mortality shocks, such as the COVID-19 pandemic, where it is crucial to incorporate future information on the expected return of mortality levels.

After its publication in 1992, the LC method became very popular, with several applications and critics, leading to the development of new variants. Lee and Miller proposed one of them in 2001 (LM) to address some of the main limitations of the LC, specifically: (i) the jump-off error, (ii) the assumption of fixed changes in the mortality pattern over time, and (iii) the adjustment of the time index \hat{k}_t [23].

In this work, we apply the LC and LM methods to forecast life expectancy at birth in São Paulo city until 2050. These models were chosen due to their simplicity and widespread acceptance in mortality forecasting. To determine a_x for the LC method, we calculated the mean of all the $\ln(m_{x,t})$ up to 2019, while for the LM, the a_x was defined as equal $\ln(m_{x,t})$ of 2019. All other estimated parameters and adjustments were made in accordance with the proposals of the authors for both methods.

Starting points There has been a significant change in living conditions in São Paulo over the last century. This has resulted in non-linear changes in the mortality pattern of the city [2]. When it comes to forecasting, historical trends are usually extrapolated into the future, including the Lee-Carter [22] and Lee-Miller [23] methods. However, this can lead to considerably different forecasted results depending on the chosen baseline period, especially if there are non-linear changes in the mortality level over time. To address this

issue, Lee and Miller [23] suggested using data after 1950 to better meet the assumption of fixed b_x . Similarly, Hyndman and Booth [27] adopted the starting year of 1950 to avoid external influences in the mortality trend, such as the two world wars and the Spanish Influenza in 1918, as well as structural changes over the last century.

The choice of using 1950 as a starting point is subjective and may not be applicable to all countries. While some studies such as those by Lee and Miller [23], Tuljapurkar et al. [28], and Hyndman and Booth [27] suggest that 1950 is a suitable starting point, it is uncertain whether this year accurately represents mortality patterns in less developed countries. For instance, in Brazil, mortality rates began to decrease after 1940 [1], and there were non-linear declines throughout the last century, resulting in varying rates of increase in life expectancy at birth [2].

This study used different starting points to forecast life expectancy until 2050, aiming to prevent the subjectivity of choosing periods with fast or slow mortality declines. The approach was to use baseline periods starting at each year between 1920 and 1995 to forecast life expectancy in São Paulo city from 2023 to 2050. For the LM method, baseline periods started between 1950 and 1995. As such, for each COVID-19 scenario and choice of the final year of the fitting period (which will be presented in the next subsection), we run the LC method for 76 different baseline periods and the LM for 46, separating men and women.

Scenarios for COVID-19 The COVID-19 pandemic led to a significant increase in deaths, particularly in Brazil, which had the second-highest number of excess deaths after COVID-19 [8]. It's uncertain whether mortality levels will return to pre-pandemic levels despite the decline in COVID-19 deaths and the widespread acceptance of vaccinations

among the population [17]. To address this issue, we propose using six different scenarios for the recovery of mortality levels in the years following COVID-19. These scenarios were computed by including or not the observed death rates of the three years of COVID-19 (2020, 2021 and 2022) and the creation of assumptions for the years after 2022.

The first scenario utilized data until 2019 and assumed that the mortality level observed before the pandemic would continue in the following years. The second scenario used data until 2022, factoring in the pandemic's impact on mortality rates over the past three years. The third scenario considered the three years of COVID-19 and a hypothetical year of 2023, assuming that mortality levels would return to pre-pandemic rates (i.e., 2019). Moreover, the fourth scenario also included the three years of COVID-19 and a hypothetical year of 2023, basing mortality levels on the forecasted levels from the first scenario.

Considering the challenges faced during the COVID-19 vaccination campaign in Brazil, which began in 2021 [16; 15], it is likely that the effects of the pandemic may only be fully realized in 2024. To explore this further, we have applied the methodology used by the United Nations in their World Population Prospects [5] to two alternative scenarios, where mortality levels return to pre-pandemic levels in 2024 instead of 2023. Scenario five assumes three years of COVID-19 (2020, 2021 and 2022), followed by a hypothetical year in 2024 with mortality levels equal to those observed in 2019, and a hypothetical year in 2023 with mortality levels between those of 2022 and 2024. In this scenario, it is assumed that mortality levels will return to pre-pandemic levels in 2024, with 2023 being an intermediary year. Scenario six also assumes three COVID-19 years (2020, 2021 and 2022), but with a hypothetical year in 2024 based on the forecasted year from scenario one (following the same starting year iteration), and a hypothetical year in 2023 with

mortality levels between those of 2022 and 2024. In this scenario, it is assumed that mortality levels will return to the expected levels for 2024 amidst the pandemic, with 2023 being an intermediary year. Finally, we recall here that for each of these six scenarios, we employ a variety of starting years for the fitting period (see previous subsection), resulting in 456 (76×6) different fitting periods for the LC method and 276 (46×6) fitting periods for the LM.

Figure 1 summarizes all six scenarios. It is important to note that other mortality outcomes may occur in São Paulo city in the next years. However, by including the six scenarios into the forecasts, it was possible to account for some of the uncertainty regarding the mortality level's return post-pandemic.

| Scenario | Data used | Assumption |
|----------|--|---|
| I | Until 2019 | We consider that the pattern observed before COVID-19 will continue to be seen until 2050. |
| II | Until 2022 | We consider that there will be a shift in mortality level due to COVID-19 that won't be absorbed. |
| III | Until 2022 2023 = 2019 | With a hypothetical year 2023 equal to 2019 we consider that mortality will return in 2023 to the same level observed in 2019. |
| IV | Until 2022 2023 = 2023 forecasted in I | With a hypothetical year 2023 equal to the forecast obtained in scenario I, we consider that mortality will return in 2023 to the level expected before COVID-19. |
| V | Until 2022 2023 = mean of 2022, 2024 2024 = 2019 | With a hypothetical year 2024 equal to 2019 we consider that mortality will return in 2024 to the same level observed in 2019. |
| VI | Until 2022 2023 = mean of 2022, 2024 2024 = 2024 forecasted in I | With a hypothetical year 2024 equal to the forecast obtained in scenario I, we consider that mortality will return in 2024 to the level expected before COVID-19. |

Figure 1 – Summary of the six scenarios considered for the return of mortality levels after the pandemic

Final estimate and prediction intervals To compute prediction intervals for life expectancy, we use a simulation approach for the forecasted k_t s. Specifically, we simulated the forecasted k_t s 100 times and computed the life expectancy at birth for each simulation over the forecasted horizon (2023-2050). This resulted in a total of 45,600 forecasts of life expectancy at birth between 2023 and 2050 for the LC method and 27,600 forecasts for the LM method for men and women separately. From these estimates, we obtained the median and the 2.5 and 97.5 quantiles. The quantiles allowed us to compute the 95% prediction interval for each forecasted year, with the median used as the final estimate.

Our final approach considered different baseline periods, six scenarios of survival recovery after the pandemic, and 100 simulations of the time-index parameter, allowing us to better account for the forecast uncertainty in the context of rapid changes in mortality levels. However, it's important to note that our analysis does not account for other potential mortality outcomes that may occur in São Paulo city in the coming years. The effect of the pandemic on mortality and other changes in mortality patterns induced by future shocks are yet to be predicted.

RESULTS

Figure 2 shows the logarithm of age-specific death rates by sex for São Paulo from 1920 to 2022. The figure reveals a consistent mortality decline over the years, with the highest level observed in 1920. Comparing the age groups, the most considerable improvement in mortality, especially among females, occurred at younger ages (up to age 50). Our

results align with previous research, which has consistently shown higher mortality rates among men than women. Also, the age patterns follow previous studies [29; 30], with decreases over time that led to a structure of higher mortality rates at birth that decreases until around age ten and increases again with a hump between ages 15-30. This peak, more pronounced among men, is attributed to external causes of death, such as accidents and violence, which have become more prevalent in Brazil over the last decades, affecting especially men at young ages [31].

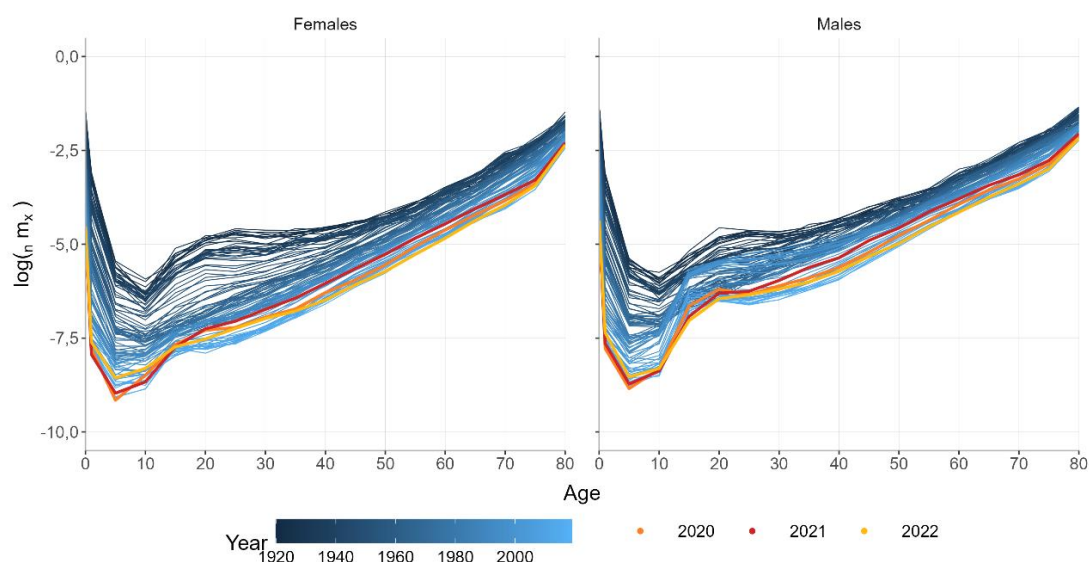


Figure 2 – Log of age-specific death rates by sex, São Paulo city 1920-2022.

Figure 2 highlights the years 2020, 2021, and 2022 in orange, red, and yellow, respectively, to indicate the period affected by COVID-19. The pandemic disrupted the pattern of decreasing death rates, causing a significant increase in mortality, especially for those above 20 years old. It is evident that the year 2021 had a more substantial impact on mortality than 2020, with a higher number of reported deaths. However, with the start

of vaccination in 2021, there was a subsequent decline in the death rates in 2022. It is highly likely that mortality will revert to pre-pandemic levels in 2023 or the following years. Regarding the age pattern of mortality, the pandemic changed the hump: the peak shifted to older age groups, reflecting variations in fatality rates across different ages, particularly the concentration of fatalities among older individuals [32]. After 2021, the hump seems to shift back to younger age groups.

Figure 3 reveals that life expectancy for both men and women has improved significantly since 1920. In that year, the level was around 40 and 43 years for men and women, respectively, whereas in 2019, it was close to 76 and 83 years. Nevertheless, over this period, life expectancy at birth did not follow a linear trend, followed by a non-expected drop after 2019 due to the COVID-19 pandemic. Between 2019 and 2021, life expectancy at birth has dropped by more than four years for males and more than three years for females. In 2022, however, life expectancy levels bounced back, returning to similar levels as of 2019. By 2022, male and female life expectancy at birth is estimated to be around 74 and 81 years old, respectively.



Figure 3 – Life expectancy at birth by sex, São Paulo city 1920-2022 and the Best-Practice slope.

São Paulo presented a complex and multifaceted mortality decline in the past century, with differences by sex and age groups. After 2000, life expectancy at birth for women in the city was already close to the estimated levels of the Best-Practice countries for men by Oeppen and Vaupel [33]. This result expresses the relatively fast convergence of mortality levels to the ones in the wealthiest countries, and it is in line with results by Vallin and Meslé [34]. By 2019, the difference between São Paulo and Best-Practice life expectancy at birth was around six years for females and seven years for males.

We analyzed mortality trends up to 2022 and used the Lee-Carter [22] and Lee-Miller [23] methods to forecast life expectancy in São Paulo city until 2050. Our methodology which included accounting for different baseline periods and scenarios for the return of mortality levels after COVID-19, led to 73,200 estimates of life expectancy at birth for

each year between 2023 and 2050. We combined the results into median and 95% prediction intervals for both males and females.

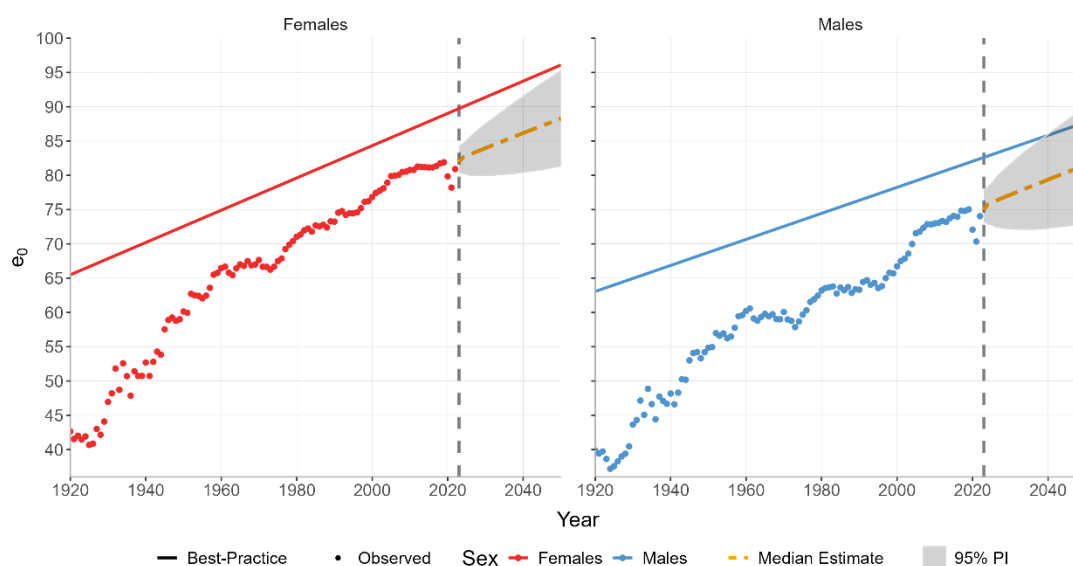


Figure 4 – Life expectancy at birth forecasts with 95% prediction intervals until 2050, São Paulo city.

Figure 4 shows life expectancy forecasts by sex. The results for females are displayed in the left panel. The median forecast for 2023 is 81.92, with a 95% PI ranging from 80.46 to 84.21. The estimated increase for this year is enough to reach the pre-pandemic level. For 2024, the median estimate is 82.59, and the 95% PI ranges from 80.29 to 84.56. From 2024 onwards, there is a steady increase in life expectancy until 2050, with estimated life expectancy at birth equal to 88.29 and 95% PI between 81.36 and 95.24.

The prediction intervals for females are relatively narrow, indicating more undersized uncertainty in the forecasts. This is likely due to the more consistent decrease in mortality

over the last century and the shorter impact of COVID-19 on female mortality. The widest 95% prediction interval is estimated in 2050, with a range of approximately 14 years.

The right panel shows forecasts for male life expectancies. The median forecast for 2023 is 74.96, with a 95% PI between 73.31 and 77.76. Although this is higher than the life expectancy observed in 2022 it is somewhat below the pre-pandemic level. For 2024, the median estimate of 75,74 is slightly above the level seen in 2019, with a 95% PI ranging from 72.95 to 78.25. From 2024 onwards, the estimates show a relatively stable increase until 2050, with estimated life expectancy at birth of 81.5 and a 95% PI ranging from 73.03 to 90.

It is essential to note that the prediction intervals are quite broad compared to females. This expresses the uncertainty surrounding the life expectancy levels when the different baseline periods and levels of mortality post-pandemic are accounted for in the forecasts. Additionally, it reflects the more disruptive impact of the pandemic on male mortality levels. Depending on the chosen baseline period and COVID-19 scenario, results with differences of up to 17 years in 2050 can be found within the 95% range.

It should be noted here that our forecasts – despite the occurrence of COVID-19 – are significantly higher than those that have been made in the past. Specifically, we forecast that life expectancy at birth in 2050 for males and females will be 2.3 and 3.59 years higher, respectively, than what was predicted by the most recent forecast in 2013 [10].

Also, within the 95% PI of our forecasts, both male and female São Paulo life expectancy could be reaching the Oeppen-Vaupel Best-Practice [33] after 2050. For this analysis, we assumed that the Best-Practice increase will follow the same pace of increase until 2050.

Additional file 1 summarizes all median estimates and 95% prediction intervals for males and females between 2023 and 2050.

DISCUSSION

The aim of this work was to forecast life expectancy at birth in São Paulo city until 2050 while considering the impact of the COVID-19 pandemic and the non-linear changes in mortality trends that have occurred over the past century. São Paulo has played a pioneering role in managing the pandemic in Brazil, and it also has a comprehensive historical record of age-specific mortality rates dating back to 1920. Therefore, using death rates for São Paulo was ideal for applying forecasting methods to analyze future survival levels in Brazil.

We used the LC and LM methods with death rates obtained from the combination of deaths gathered by the SEADE foundation (SEADE) and population collected by the Brazilian Institute of Geography and Statistics (IBGE) data. These methodologies have gained popularity among demographers due to their innovative and straightforward approach to mortality forecasting, although they are not without limitations. Most forecasting methods, including the LC and the LM, are extrapolative. Therefore, these methods are sensitive to the choice of the baseline period and fluctuations in the rate of mortality decline in the past.

As emphasized in our study, from 1920 to 2019, life expectancy at birth in São Paulo increased non-linearly. Moreover, there was a substantial reduction in life expectancy during the years of the pandemic. To mitigate the dependency on the fitting period's

choice and better incorporate the effects of the recent mortality shock caused by COVID-19, we used different baseline periods, from 1920 and 1995 until 2022, and six scenarios for post-pandemic mortality levels. Also, we simulated the forecasted time-index parameters to compute prediction intervals.

Based on 73,200 simulations for each year between 2023 and 2050, we summarized the resulting life expectancy forecasts into median values and 95% prediction intervals. These outcomes improved the accuracy of life expectancy forecasts for São Paulo. We are unaware of any estimates that have accounted for pandemic effects and explored the non-linear changes in mortality patterns over the past century in São Paulo. Also, adding recent data enhanced the forecasts by considering the city's most recent mortality experiences. Despite the negative effects of the COVID-19 pandemic, our results presented significantly higher estimates than predicted by the most recent forecast in 2013 [10]. By 2050, we estimate that life expectancy at birth in São Paulo city will reach approximately 81.5 (95% PI ranging from 73.03 to 90) for men and 88.29 (95% PI ranging from 81.36 to 95.24) for women. Also, within the 95% PI, we estimated that after 2050, male and female life expectancy could reach the Best-Practice levels established by Oeppen and Vaupel [33], extrapolated for the future.

Our forecasting approach – based on a combination of different fitting periods and future scenarios – is among the first attempts to forecast mortality in the presence of shocks. We first proposed this approach in 2023 [35], and our forecast for 2022 was extremely close to the subsequent observed data – 80.9 years observed versus 80.9 years forecasted for females and 74 years observed versus 73.6 years forecasted for males. However, the long-term effects of the pandemic on health and mortality remain unclear. Although vaccination rates are increasing in São Paulo city, the worsening of the interactions of

COVID-19 with other illnesses, such as influenza, or the occurrence of new variants is yet to be predicted. Therefore, it is essential to note that different mortality outcomes - not accounted for in our six scenarios - may occur in the city in the coming years.

Finally, by including different baseline periods, starting from 1920 to 1995 until 2022, we were able to combine forecasts obtained over different estimated b_x s, one for each baseline period, and consequently, account for some possible changes in the age pattern over time. Future research could, however, focus on applying methods that account for variations in the b_x , such as Li, Lee, and Gerland [24], or methodologies that are not based on the Lee-Carter framework.

CONCLUSIONS

In conclusion, this study provides a comprehensive analysis of future life expectancy in São Paulo city, considering the significant impact of COVID-19 and the non-linear changes in mortality trends observed over the past century. Using different baseline periods and post-pandemic scenarios allows for a more nuanced understanding of the potential trajectories of life expectancy, reflecting both historical and recent mortality patterns. Our findings indicate a substantial recovery in life expectancy post-pandemic in the short-term, and significant increases for both men and women by 2050.

The analysis and forecasts presented in this study provide a critical resource for policymakers and researchers focused on addressing public health challenges arising from the pandemic and improving population well-being. However, our results share the same

limitations as other extrapolation methods. Historical patterns may not hold in the future, and unpredicted structural changes, such as new mortality shocks, may occur.

ABBREVIATIONS

IBGE – Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatística)

HMD – Human Mortality Database

LC – Lee-Carter method

LM – Lee-Miller method

SEADE – SEADE foundation (Fundação SEADE)

Availability of data and materials

Data is available upon request. R codes are available at the open-access GitHub repository https://github.com/marialaura-mlm/ForeCOVID_SP.

Authors' contributions

All authors conceptualized the study and designed the methodology. MM wrote the first draft of the manuscript, carried out the visualizations, and performed the formal analysis. MM and UB contributed to the programming and software development. UB and CT provided supervision. All authors reviewed and edited the manuscript.

Acknowledgments

The authors wish to thank Bernardo Queiroz and Enrique Acosta for the feedback given on a previous version of this paper and Larissa Souza for sharing the data with corrections until 2000.

REFERENCES

1. Prata PR. A Transição Epidemiológica no Brasil. *Cad Saude Publica*. 1992;8(2):168–75.
2. Simões CCS. Perfis de saúde e de mortalidade no Brasil: uma análise de seus condicionantes em grupos populacionais específicos. Brasília, 1ª edição. 2001.
3. Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proc Natl Acad Sci USA*. 2020;117:9696–8.
4. Nepomuceno MR, Acosta E, Alburez-Gutierrez D, Aburto JM, Gagnon A, Turra CM. Besides population age structure, health and other demographic factors can contribute to understanding the COVID-19 burden. *Proc Natl Acad Sci USA*. 2020;117(25):13881–3. <https://doi.org/10.1073/pnas.2008760117>
5. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2022*. Available in: <https://population.un.org/wpp/>. Accessed May 30 2023.
6. Schöley J, Aburto JM, Kashnitsky I, et al. Life expectancy changes since COVID-19. *Nat Hum Behav*. 2022. <https://doi.org/10.1038/s41562-022-01450-3>
7. Castro MC, Gurzenda S, Turra CM, et al. Reduction in life expectancy in Brazil after COVID-19. *Nat Med*. 2021;27:1629–35. <https://doi.org/10.1038/s41591-021-01437-z>
8. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. *eLife*. 2021;10.
9. Marinho F, et al. Excess mortality in Brazil: a detailed description of trends in mortality during the COVID-19 pandemic. *Vital Strategies*. 2020. Available from: <https://www.vitalstrategies.org/resources/excess-mortality-in-brazil-a-detailed-descriptionof-trends-in-mortality-during-the-covid-19-pandemic/>

10. Fundação SEADE (SEADE). 459 anos de São Paulo Retrato da população hoje e as necessidades rumo aos 500 anos. Resenha de Estatísticas Vitais do Estado de São Paulo Ano 13 – nº 1. Available in: <https://economia.seade.gov.br/wp-content/uploads/sites/18/2021/12/SPDemografico-459-anos-sao-paulo-retrato-populacao-hoje-necessidades-rumo-500-anos.pdf>. Accessed 30 May 2023.
11. Riffe T, Acosta E, the COVERAGE-DB team. Data Resource Profile: COVERAGE-DB: a global demographic database of COVID-19 cases and deaths. *Int J Epidemiol.* 2021;50(2):390–390f. <https://doi.org/10.1093/ije/dyab027>
12. Castro MC, Gurzenda S, Turra CM, et al. Research Note: COVID-19 Is Not an Independent Cause of Death. *Demography.* 2023;60(2):343–9. <https://doi.org/10.1215/00703370-10575276>
13. Schwartz J. Harvesting and long-term exposure effects in the relation between air pollution and mortality. *Am J Epidemiol.* 2000;151:440–8.
14. Dey S, Davidson J. The determinants of non-COVID-19 excess deaths during the COVID-19 pandemic: A cross-country panel study. *Stud Microecon.* 2021;9:196–226.
15. Barberia LG, Gómez EJ. Political and institutional perils of Brazil's COVID-19 crisis. *Lancet.* 2020;396:367–8.
16. Castro R. Vacinas contra a Covid-19: o fim da pandemia?. *Rev Saude Coletiva Rio de Janeiro.* 2021;31(1).
17. Ministério da Saúde (Ministry of Health). Programa Nacional de Imunizações. Available in: <https://www.gov.br/saude/pt-br/coronavirus/vacinas/pni>. Accessed May 26 2023.
18. Conselho Nacional de Secretários de Saúde (CONASS). Painel Conass Covid-19. Available in: <https://www.conass.org.br/painelconasscovid19/>. Accessed 26 May 2023.
19. Souza LG. Diferencial de gênero na mortalidade no município de São Paulo, 1920 a 2020: padrões por idade e causas de morte. [Unpublished dissertation]. Federal University of Minas Gerais; 2023.

20. Silva FS. Tendências e projeção da mortalidade do município de São Paulo – 1920 a 2100. [Unpublished master's thesis]. Federal University of Minas Gerais; 2009.
21. Siviero PCL. Níveis e padrões do diferencial de mortalidade por sexo no município de São Paulo, 1920 – 2005. [Unpublished master's thesis]. Federal University of Minas Gerais; 2009.
22. Lee RD, Carter LR. Modeling and Forecasting U.S. Mortality [with Rejoinder]. *J Am Stat Assoc.* 1992;87(419):659–71.
23. Lee RD, Miller T. Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality. *Demography.* 2001;38:537–49.
24. Li N, Lee R, Gerland P. Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections. *Demography.* 2013;50(6):2037-51. doi: 10.1007/s13524-013-0232-2.
25. Basellini U, Camarda CG, Booth H. Thirty years on: A review of the Lee–Carter method for forecasting mortality. *Int J Forecast.* 2023;39(3):1033-49. <https://doi.org/10.1016/j.ijforecast.2022.11.002>.
26. Lee R. The Lee-Carter method for forecasting mortality, with various extensions and applications. *N Am Actuar J.* 2000;4(1):80-91.
27. Hyndman RJ, Booth H. Stochastic population forecasts using functional data models for mortality, fertility and migration. *Int J Forecast.* 2008;24(3):323–42.
28. Tuljapurkar S, Li N, Boe C. A universal pattern of mortality decline in the G7 countries. *Nature.* 2000;405(6788):789–92.
29. Ebeling M. How Has the Lower Boundary of Human Mortality Evolved, and Has It Already Stopped Decreasing?. *Demography.* 2018;55:1887–903. <https://doi.org/10.1007/s13524-018-0698-z>
30. Wilmoth JR. In Search of Limits: What Do Demographic Trends Suggest about the Future of Human Longevity. In: Wachter KW, Finch CE, editors. *Between Zeus and the salmon:*

- The biodemography of longevity. Washington, DC: National Academy Press; 1997. p. 38-64.
31. Borges GM. Health transition in Brazil: Regional variations and divergence/convergence in mortality. *Cad Saude Publica*. 2017;33(8):1–15.
 32. Levin AT, Hanage WP, Owusu-Boaitey N, et al. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol*. 2020;35:1123–38. <https://doi.org/10.1007/s10654-020-00698-12>
 33. Oeppen J, Vaupel J. Broken limits to life expectancy. *Science*. 2002;296:1029–31. <https://doi.org/10.1126/science.1069675>
 34. Vallin J, Mesle F. Convergences and divergences in mortality: A new approach to health transition. *Demogr Res*. 2004;10(2):11–44.
 35. Miranda et al. Forecasting Life Expectancy in São Paulo City, Brazil, amidst the COVID-19 Pandemic. [Unpublished master's thesis]. Federal University of Minas Gerais; 2009. Available from: <https://repositorio.ufmg.br/handle/1843/60064>

ANNEXES

Annex 1 – Summary of life expectancy estimates and 95% prediction intervals by sex, São Paulo city 2023-2050

| Year | Males | | | Females | | |
|------|--------|----------------|----------------|---------|----------------|----------------|
| | Median | 95% Lower bond | 95% Upper bond | Median | 95% Lower bond | 95% Upper bond |
| 2023 | 74.96 | 73.31 | 77.76 | 81.92 | 80.46 | 84.21 |
| 2024 | 75.74 | 72.95 | 78.25 | 82.59 | 80.29 | 84.56 |
| 2025 | 75.99 | 72.83 | 78.72 | 82.83 | 80.18 | 85.01 |
| 2026 | 76.24 | 72.52 | 79.46 | 83.06 | 80.00 | 85.58 |
| 2027 | 76.48 | 72.41 | 80.06 | 83.29 | 80.02 | 86.08 |
| 2028 | 76.71 | 72.32 | 80.60 | 83.51 | 80.00 | 86.55 |
| 2029 | 76.93 | 72.25 | 81.11 | 83.73 | 79.99 | 87.00 |
| 2030 | 77.15 | 72.20 | 81.57 | 83.95 | 80.00 | 87.46 |
| 2031 | 77.38 | 72.18 | 82.05 | 84.17 | 80.01 | 87.86 |
| 2032 | 77.60 | 72.16 | 82.54 | 84.40 | 80.02 | 88.26 |
| 2033 | 77.82 | 72.14 | 83.02 | 84.62 | 80.08 | 88.68 |
| 2034 | 78.05 | 72.14 | 83.45 | 84.83 | 80.11 | 89.09 |
| 2035 | 78.27 | 72.14 | 83.85 | 85.05 | 80.15 | 89.49 |
| 2036 | 78.48 | 72.13 | 84.28 | 85.28 | 80.23 | 89.90 |
| 2037 | 78.71 | 72.18 | 84.73 | 85.49 | 80.25 | 90.30 |
| 2038 | 78.91 | 72.20 | 85.15 | 85.72 | 80.31 | 90.71 |
| 2039 | 79.14 | 72.22 | 85.57 | 85.94 | 80.37 | 91.10 |
| 2040 | 79.36 | 72.30 | 85.98 | 86.15 | 80.46 | 91.48 |
| 2041 | 79.58 | 72.36 | 86.37 | 86.37 | 80.55 | 91.88 |
| 2042 | 79.80 | 72.44 | 86.74 | 86.58 | 80.60 | 92.23 |
| 2043 | 80.02 | 72.47 | 87.15 | 86.80 | 80.69 | 92.65 |
| 2044 | 80.22 | 72.57 | 87.57 | 87.02 | 80.79 | 93.03 |
| 2045 | 80.45 | 72.64 | 87.98 | 87.22 | 80.82 | 93.41 |
| 2046 | 80.65 | 72.69 | 88.37 | 87.43 | 80.96 | 93.76 |
| 2047 | 80.87 | 72.76 | 88.80 | 87.65 | 81.06 | 94.11 |
| 2048 | 81.08 | 72.82 | 89.23 | 87.86 | 81.16 | 94.49 |
| 2049 | 81.29 | 72.94 | 89.65 | 88.07 | 81.27 | 94.88 |
| 2050 | 81.50 | 73.03 | 90.00 | 88.29 | 81.36 | 95.24 |