



## Original article

# Spatiotemporal ecological study of COVID-19 mortality in the city of São Paulo, Brazil: Shifting of the high mortality risk from areas with the best to those with the worst socio-economic conditions

Patricia Marques Moralejo Bermudi<sup>a</sup>, Camila Lorenz<sup>a,\*</sup>, Breno Souza de Aguiar<sup>b</sup>, Marcelo Antunes Failla<sup>b</sup>, Ligia Vizeu Barrozo<sup>c,1</sup>, Francisco Chiaravalloti-Neto<sup>a,1</sup>

<sup>a</sup> Departamento de Epidemiologia, Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo, SP, Brazil

<sup>b</sup> Gerência de Geoprocessamento e Informações Socioambientais (GISA), da Coordenação de Epidemiologia e Informação (CEInfo) da Secretaria Municipal de Saúde de São Paulo, SP, Brazil

<sup>c</sup> Departamento de Geografia, Faculdade de Filosofia, Letras e Ciências Humanas e Instituto de Estudos Avançados da Universidade de São Paulo, São Paulo, SP, Brazil

## ARTICLE INFO

## Keywords:

Health inequity  
Integrated nested laplace approximation  
Spatio-temporal analysis  
Mortality  
Pandemics

## ABSTRACT

**Background:** Currently, Brazil is experiencing one of the fastest increasing coronavirus disease (COVID-19) mortality rates worldwide, with a minimum of 158,000 confirmed deaths presently. The city of São Paulo is particularly vulnerable because it is the most populated city in Brazil. Thus, this study aimed to analyse COVID-19 mortality in a spatiotemporal context in São Paulo, with respect to socio-economic levels.

**Method:** We modelled the deaths using spatiotemporal architectures and Poisson probability distributions using a latent Gaussian Bayesian model approach.

**Results:** Both total deaths and confirmed deaths showed similar spatial patterns. Mortality was higher in men and increased with age. The most critical period regarding mortality occurred between the 20th and 23rd epidemiological weeks, followed by an apparent stabilisation of the epidemiological trend. The risk of death was greater in areas with the worst social conditions during the study period. However, this pattern was not uniform over time, since we identified a shift of high risk from the areas with the best socio-economic conditions to those with the worst conditions.

**Conclusions:** Our study corroborated the relationship between COVID-19 mortality and socio-economic conditions, revealing the importance of geographic screening in the integration of better actions to face the pandemic.

## 1. Introduction

Mortality owing to coronavirus disease (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is considered avoidable because some collective and individual measures can help prevent the infection, and appropriate health assistance greatly reduces the mortality risk. Nevertheless, mortality risk varies according to some individual and geographic risk factors, resulting in health inequity observed in several parts worldwide since the beginning of the pandemic [1,2]. In Brazil, COVID-19 was first reported in the city of São Paulo on February 25, 2020 [3]. Until 3 August, 4 months after the first reported death, the disease had already resulted in 2,962,442 confirmed cases and 99,572 deaths in the country [4]. Currently, the United States

of America and Brazil are the epicentres of the disease.

Brazil is the fifth country worldwide in terms of surface area and population and is classified as having an upper-middle-income economy [5]. Almost 60% of the Brazilian population is concentrated in 6% of the large cities, among which São Paulo is the largest. The Brazilian Unified Health System guarantees healthcare for all citizens as well as for thousands of foreigners residing or passing through the country [6]. Despite this universal healthcare, geographic differences in mortality rates across different areas have been observed on a national and intra-urban scale. Spatial heterogeneity in population characteristics such as age, underlying health, household densities, partial lack of sanitation, socio-economic status, contact networks, and mobility patterns [7] has emerged as a potential propellant of the spatiotemporal

\* Corresponding author. Department of Epidemiology, School of Public Health, University of Sao Paulo, Brazil.

E-mail address: [camilalorenz@usp.br](mailto:camilalorenz@usp.br) (C. Lorenz).

<sup>1</sup> These senior authors contributed equally.

spread of the disease.

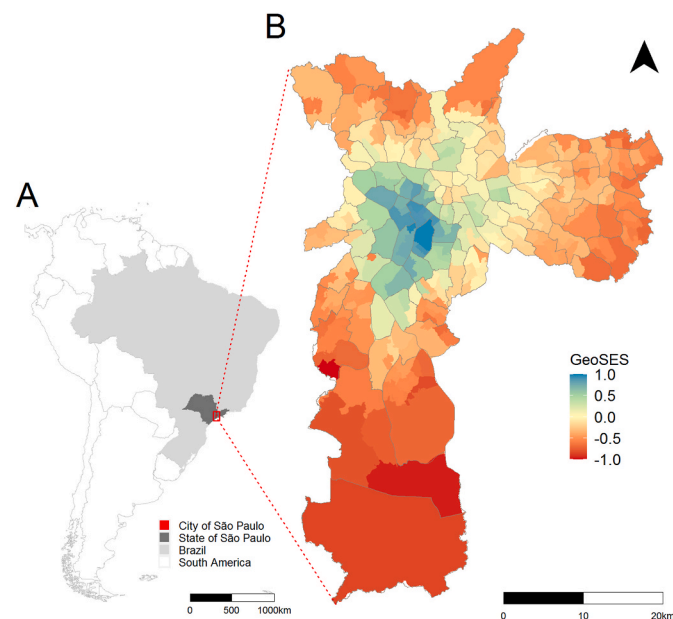
Since the beginning of the pandemic, mapping disease occurrence and spread has become a powerful tool to track the disease and establish measures to slow down the transmission of the infection both locally and globally [8]. Web-based Geographical Information Systems have allowed near real-time monitoring using map-centric dashboards [9–11]. Despite advances in the use of technology to reduce the impact of the pandemic, little is known about the spatiotemporal patterns of COVID-19 mortality, especially in intra-urban settings. Studying the spatiotemporal dynamics of mortality instead of the cases may help better evaluate inequity. All health disadvantages accumulated over decades of life, owing to any kind of deprivation, increase the risk of mortality due to COVID-19. The lack of a robust spatiotemporal analysis undermines the comprehension of mitigation strategies to potentialise disease-control efforts. Thus, this study aimed to unveil the spatiotemporal dynamics of COVID-19 mortality in a higher spatial resolution (in the city of São Paulo), considering the socio-economic context of the population. This approach may shed light on the urgent need for solid evidence on health inequities during the COVID-19 outbreak.

## 2. Materials and methods

### 2.1. Study area and data acquisition

This ecological study, based on COVID-19 secondary mortality data, was delineated in the city of São Paulo, state of São Paulo, Brazil (Fig. 1). In 2020, the estimated population of this city was 11,869,660 inhabitants, and the mean demographic density was 7803 inhabitants/km<sup>2</sup> [12].

When using COVID-19 data, underreporting is always an issue that deserves attention, even for mortality. To minimise the effect of possible sub-notification of deaths, we analysed confirmed and suspected deaths due to COVID-19. Thus, mortality data comprised confirmed and suspected deaths that occurred between 15 March and June 13, 2020, extracted on 18 June from the Mortality Information System (SIM), the Mortality Information Improvement Program (PRO-AIM) of the Epidemiology and Information Coordination (CEInfo) of the São Paulo Health Secretariat (SMS-SP). Confirmed deaths due to COVID-19 corresponded



**Fig. 1.** a) South America, Brazil, State of São Paulo, city of São Paulo. b) Distribution of socio-economic Index of the Geographic Context for Health Studies (GeoSES) according to the sample area and delimited by administrative districts, city of São Paulo, 2010.

to the code B34.2 (coronavirus infection disease), according to the International Classification of Diseases Tenth Revision (ICD-10). Suspected deaths were coded as U04.9 (corresponding to severe acute respiratory syndrome).

The places of residence of the COVID-19 deaths were geocoded with CEInfo/SMS-SP using its databases and Google Maps API geocoding script (that uses public places as the base map). The resulting geocoded addresses were validated by comparing the road or Zone Improvement Plan (ZIP code), whenever the record was allocated using the original ZIP code. Geocoded data were assigned to the 310 sample areas of the Brazilian Institute of Geography and Statistics, for which demographic and socio-economic census data are available [13]. We considered these areas as the spatial units in our models (Fig. 1).

Records, including the basic cause of death, age, sex, date of death according to the epidemiological week (EW) [14], and the sample areas of residence, were obtained after addressing a formal request to the São Paulo Electronic Information System (e-SIC database, protocol 48567). Their data information is hosted in an open session on the municipality's transparency portal for public access [15]. Here, we named this information the e-SIC database. It was not necessary to submit this study to an ethics committee because we did not have access to personal data, such as names and addresses. The use of secondary data without personal identification and in a public domain dispenses the need for prior approval from the Ethics Committee on Research with Human Beings (as per Resolution No. 510/2016 of the National Health Council) [16].

We also used the data available in the TabNet and named this information as the "Tabnet database". The Tabnet database is an application, available at <<https://www.prefeitura.sp.gov.br/cidade/secretarias/saude/tabnet/>>, provided by the Municipal Health Department of São Paulo. It was developed by DataSUS. This application provides free access (to any user) to population databases and to database information systems of SUS (such as the SIM), which is supplied by the Secretariat's Program (PRO-AIM). Through the Tabnet database application, it is possible to perform tabulations and to cross several variables of interest such as EW, sex, age group, and specific cause. The databases are updated periodically. Notably, the data sources of the Tabnet and e-SIC database are SIM. However, only the e-SIC database is subdivided according to sample area.

To measure the socio-economic condition of the population, since individual-level data are not available in the mortality database, we used a socio-economic index elaborated for health research. The Socio-economic Index of the Geographic Context for Health Studies (GeoSES) [17] was developed using principal component analysis, starting with 41 variables. The index conceives the socio-economic condition by considering seven parameters based on the theoretical background [18, 19]: education, mobility, poverty, wealth, income, segregation, and deprivation of resources and services. The index was defined on three scales: national, Federative Unit, and intra-municipal. Fig. 1 presents GeoSES for the sample areas of the city of São Paulo. It shows that the areas with the best socio-economic conditions (GeoSES equal to or close to 1) are located in the central part of the city and that the socio-economic conditions deteriorate towards the periphery, where they reach the worst levels (GeoSES equal to or close to -1). It has been shown to be useful in studies of mortality due to avoidable causes of deaths in individuals aged 5–74 years due to interventions at the Brazilian health system on a national scale as well as mortality due to circulatory system diseases in the city of São Paulo [17]. This index is publicly available at <https://opendatasus.saude.gov.br/dataset/geoses>.

### 2.2. Data analysis

We used the information of confirmed (B34.2) and suspected (U04.9) deaths due to COVID-19 available for the entire city of São Paulo from the 11th to 29th EWs to calculate the weekly mortality rates of confirmed (B34.2), suspected (U04.9), and total (B34.2 + U04.9) COVID-19 deaths. This was similarly done using the e-SIC database from

the 11th to 12th EW. We excluded (from the information e-SIC database) COVID-19 deaths that occurred in the 25th EW because data from this week were incomplete (it was extracted on June 18, 2020, and it contained only part of the 25th EW information). These rates were obtained by dividing the respective numbers of deaths in each week by the total population of the city and presented as death per 100,000 inhabitants per week. In this sequence, we obtained the mortality rates for confirmed, suspected, and total COVID-19 deaths according to sex and age for the entire period from the 11th to 12th EW using the Tabnet and e-SIC databases. The comparisons between the data from these two sources were useful to evaluate the completeness of the data we used for the spatial and spatiotemporal analysis. For the calculation of the mortality rates according to sex and age, we excluded the data without these information.

As only one suspected COVID-19 death occurred during the 11th EW, we restricted our spatial and spatiotemporal analysis from the 12th to 24th EW, and spatial or spatiotemporal architecture was considered in all models developed. We first modelled the confirmed and total COVID-19 deaths using spatiotemporal models involving only the intercept and random effects that account for spatial and temporal autocorrelation and the interaction between them. The spatial dependence was modelled considering the Besag–York–Mollié (BYM) model with two components representing the spatially structured and non-structured random effects [20,21]. These two components were considered independent from one another and followed the parameterisation proposed previously [22]. The temporal dependence was modelled using a non-structured random effect and a structured random effect provided by a random walk autoregressive model of first order (RW1). The interaction between space and time was modelled considering spatial and temporal non-structured random effects [21].

The number of confirmed and total COVID-19 deaths per EW and per sample area was modelled using Poisson and zero-inflated Poisson probability distributions with a latent Gaussian Bayesian model approach. We considered the expected confirmed and total COVID-19 deaths for each EW and for each spatial unit as offsets in these models. The expected deaths were estimated with indirect standardisation, considering the age and sex structure of each sample area and the mortality rates for the entire study period and city. This enabled us to interpret the outcomes of our analysis as relative risks (RR) concerning the mortality rates for the entire study period and city. From these models, we obtained the temporal and spatiotemporal RR. Subsequently, we introduced the socio-economic covariate (GeoSES) in these models and obtained the corresponding RR.

Finally, we used a spatial approach to model the confirmed and total COVID-19 death per EW to evaluate the role of the socio-economic covariate in each EW. To this end, we considered spatial models with intercept, BYM spatial random effects, and GeoSES as a covariate. The expected COVID-19 deaths were obtained in a similar way for the spatiotemporal models. We considered the entire city's mortality rates for each EW, and as such, we could interpret the RR of the entire city's mortality rates for each EW.

We performed our models in a Bayesian context using the integrated nested Laplace approximation (INLA) approach [23]. We selected the best models using the Deviance Information Criterion (DIC) so that the best-adjusted models were those with lower DIC values [21]. We used non-informative priors for the fixed effects and priors with penalised complexity for the precision parameters of the random effects [22]. We ran our models in the R environment [24].

### 3. Results

We found 14,753 confirmed and suspected COVID-19 deaths in the Tabnet database (from EW 11 to 29) and 10,760 in the e-SIC database (from EW 11 to 25) in the city of São Paulo. We excluded 67 deaths from the e-SIC database because they were referent to the 25th EW, which was not completed when the data were extracted. Accordingly, we were

left with 10,693 deaths. Fig. 2 shows the mortality rates for COVID-19 for both data sources considering the confirmed, suspected, and total deaths. The curves from the e-SIC database are similar to the curves from the Tabnet database. Both curves have the same data source with different extraction dates, and the differences among these curves in the 23rd and 24th EW are related to a delay in the notification of the COVID-19 deaths.

Table 1 shows the numbers and mortality rates of confirmed, suspected, and total COVID-19 deaths in the different sex and age groups, obtained from the e-SIC and Tabnet databases (from EW 11 to 24). To build Table 1, we excluded six deaths with ignored age and three deaths with ignored sex from the 10,693 deaths in the e-SIC database. We excluded six deaths with ignored age and four deaths with ignored sex from the 11,098 deaths in the Tabnet database. We can observe that the mortality was higher in men and that it increased with age, resulting in a mortality rate of 460.9 deaths per 100,000 inhabitants (in 14 weeks) for people aged  $\geq 60$  years. This pattern of increased mortality with increase in age was also observed after making a greater stratification among people aged 60 years or above, as shown in Supplementary Material 1.

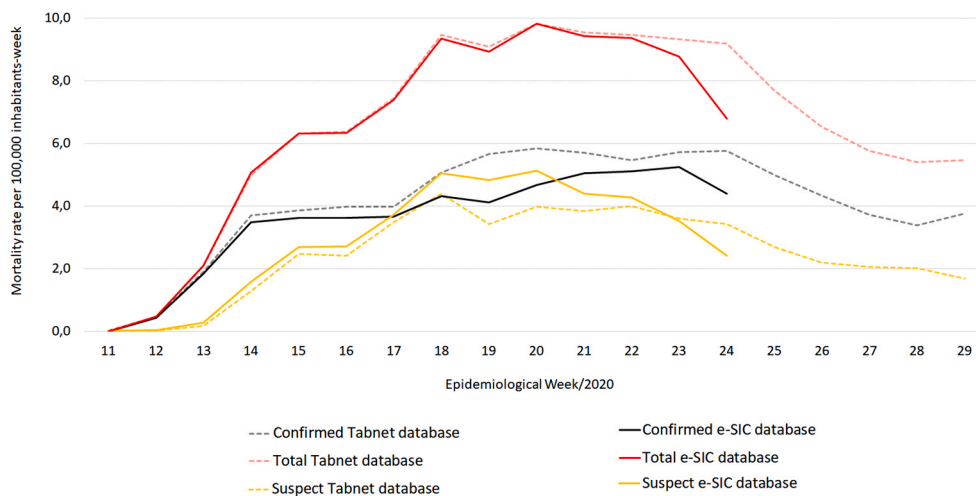
Table 1 also shows that the data we used to build our spatial models (e-SIC database) are very close to the municipality's official data on the pandemic (Tabnet database). To construct these models, we excluded suspected COVID-19 deaths that occurred in the 11th EW (the first one in the city) from the e-SIC database. We also excluded nine deaths with ignored sex or age, and 68 cases whose addresses were not geocoded or did not have the sample area codes. We achieved a high geocoding success rate of 99.3%. Out of a total of 10,692 deaths (excluding the first death in the 11th EW), 10,619 records were geocoded using the address data. Of these, 5837 referred to ICD B34.2 (99.4% of the initial total of 5875 records) and 4782 to ICD U04.9 (99.2% of the total of 4817 records).

During this period, 30.6% and 17.4% of the sample areas had zero confirmed and zero total COVID-19 deaths, respectively. The Deviance Information Criterion (DIC) values of the spatiotemporal models with Poisson probability distribution were lower than the value of the zero-inflated Poisson distribution (Supplementary Material 2). Considering each week separately, the number of zero deaths in the sample areas varied from 6.1% to 85.5%. The DIC values for the spatial models with Poisson probability distribution were, in most cases, lower than the values with zero-inflated Poisson distribution. In cases where this did not occur, they were very close to each other (Supplementary material 3). From these results, we considered the best-adjusted spatial and spatiotemporal models to be those with the Poisson distribution.

We present the results of the models with spatiotemporal architecture using only the intercept. Fig. 3 shows the temporal RR from the 12th to 24th EW. Notable features include the maximum RR, which occurred during the 20th EW for the total COVID-19 deaths and during the 23rd EW for the confirmed ones during our study period. These results, even adjusted for the temporal autocorrelation, are similar to those presented in Fig. 2. Considering the data presented in Figs. 2 and 3, the apparent pattern of the temporal curves shows a tendency to stabilise. However, a new rise in mortality cannot be ruled out.

Figs. 4 and 5 show the posterior means of the spatiotemporal RR for the sample areas and EW, respectively, for confirmed and total COVID-19 deaths. Apart from the fact that the RR is greater for the total deaths than for the confirmed ones, the distribution of the RR is similar between them, and it follows the behaviour of the temporal RR. In the first two EWs, the sample areas presented lower values of RR that increased over time. However, this increase occurred with greater intensity in peripheral areas.

We also used the spatiotemporal models, considering socio-economic variables (GeoSES). Table 2 shows the spatiotemporal RR and the 95% credibility intervals for GeoSES obtained for the models with confirmed and total COVID-19 deaths. In both models, it is noted that a high socio-economic level protected against the risk of mortality due to COVID-19



**Fig. 2.** Distribution of mortality rates (per 100,000 inhabitants-week) of suspected (U04.9), confirmed (B34.2), and total (U04.9 + B34.2) COVID-19 deaths, according to e-SIC and Tabnet databases and epidemiological week. City of São Paulo, 2020. Data source: Deaths: Mortality Information System of Mortality Information Improvement Program, from the Epidemiology and Information Coordination of the São Paulo Health Secretariat (SIM/PRO-AIM/CEInfo/SMS-SP). Population: Fundação Sistema Estadual de Análise de Dados Estatísticos (SEADE). Tabnet database was update on 7/23/2020 and e-SIC database was provided on August 6, 2020.

throughout the study period. Thus, an increase of one unit in the socio-economic indicator represented a 25% reduction in the risk of mortality due to COVID-19 (for the model using only confirmed deaths) and a 33% reduction in the risk of dying for the model using total deaths of COVID-19. Moreover, the risk of mortality due to COVID-19 in the sample areas with the best socio-economic conditions (GeoSES close to 1), compared with that in areas with the worst conditions (GeoSES close to  $-1$ ), was 50% lower for the model with confirmed deaths and 66% lower for the model with total deaths.

Finally, we performed the spatial modelling of confirmed and total COVID-19 deaths in the different EWs separately, considering the socio-economic covariate. Fig. 6 shows the RR and 95% credibility intervals for GeoSES, with respect to the confirmed and total COVID-19 deaths for each of the EW. We identified a shift in the pattern of the relationship between COVID-19 mortality and socio-economic status over time. The best socio-economic level presented itself as a risk factor for COVID-19 deaths in the first two EWs in the city of São Paulo. From the 15th EW, for total deaths, and from the 16th EW, for confirmed and total deaths, the worst socio-economic condition became a risk factor. Even if some values were not significant, there was a continuous decrease in RR from the 12th to 17th EW, followed by its stabilisation.

#### 4. Discussion

This is the first population-based study on the evolution of the spatiotemporal pattern of COVID-19 mortality in the intra-urban setting of the largest city of Brazil. Using two different datasets, analysing the confirmed and confirmed plus suspected deaths separately, we were able to evaluate how uncertainty would impact the association between the RR and the socio-economic condition. The robust EW models clearly showed when the high risk of death shifted from areas with the best to those with the worst socio-economic conditions in the city.

Our findings showed that the most critical period regarding mortality due to COVID-19 in the city of São Paulo occurred between the 20th and 23rd EW, followed by an apparent stabilisation of the temporal trend. However, it is not possible to predict a future scenario. As social distancing measures have been lifted in the city since the 25th week [25], this could increase the number of infected people and, consequently, the number of deaths. Although social distancing alone seems insufficient to contain COVID-19, many studies have frequently concluded that it remains a critical component of the outbreak control [26]. It is important to point out that both total deaths and confirmed deaths showed similar spatial patterns in our study, despite their differences. The suspected deaths, on the one hand, need to be treated with caution because they may not be COVID-19-related and, on the other hand, could be considered as one of the strengths of the study. This is

because there is a delay in the confirmation of suspected cases, and consequently, some of the suspected deaths are later confirmed to be COVID-19-related deaths. Furthermore, some of the suspected cases would be considered as confirmed COVID-19-related deaths if the case definition was not strict (as the required test is often not available or is not performed within the appropriate window time). From this perspective, the amplitudes of variation in rates and RR obtained from confirmed and total deaths could be considered as lower and upper limits (or vice versa) for the magnitudes of these measures.

The elderly population represents one of the groups that are more prone to COVID-19 in the city of São Paulo. The risk of death was the highest among men aged  $>70$  years; this is similar to data from China and the United States [27]. Recently, Souza et al. [3] analysed the Brazilian population and found that most COVID-19 deaths were of men and that the most frequent comorbidities were cardiovascular disease and diabetes. Behavioural factors, especially social status, that may prejudice adherence to lockdown measures, have been shown to be potentially crucial in determining susceptibility to SARS-CoV-2 [28,29]. This unequal death ratio in men may be interpreted considering many factors: the comparatively higher prevalence of comorbidities (i.e., hypertension, diabetes, cardiovascular disease, and chronic lung disease) [30], higher risk behaviours (i.e., smoking and alcohol use), occupational exposure [31], and sex differences in immune responses [32]. However, there may be other social and behavioural characteristics that favour women as reported in previous studies, which proposed that women are more likely than men to adopt hand hygiene practices [33] and to seek preventive care [34].

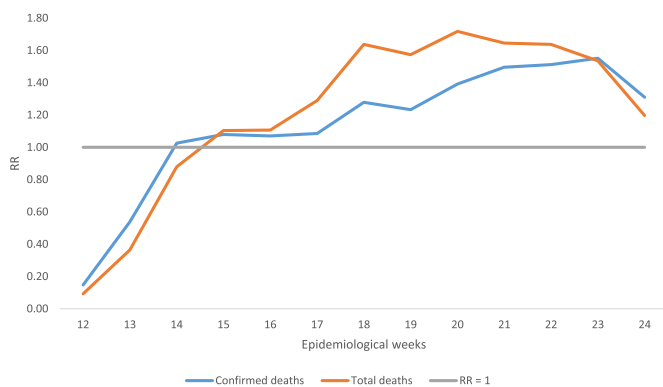
The spatial distribution of suspected and confirmed deaths due to COVID-19 in the city of São Paulo shows inequalities, with spatial dependence and positive correlation associated with socio-economic factors of the areas. This is remarkably similar to the results of Maciel et al. [35]. Our findings reveal that the socio-economic condition acts as a protective factor against the risk of mortality due to COVID-19. In the models considering only confirmed deaths and all deaths, an increase of one unit in the socio-economic indicator represented a 25% and 33% decrease in the risk of mortality, respectively. The first observation is that, when considering all deaths, the protective effect of the socio-economic level is more evident, showing that there must be a higher incidence of suspected deaths in the less favoured areas than in the most favoured areas (in areas with better socio-economic level, confirmation tests for COVID-19 are probably more available). A study conducted by Souza et al. [3] reinforced this finding. They compared the spatial pattern of confirmed cases of COVID-19 and severe acute respiratory infection with unknown aetiology, with per capita income in the metropolitan region of São Paulo. They found that the COVID-19 cases were more associated with better levels than the latter. They pointed out

**Table 1**

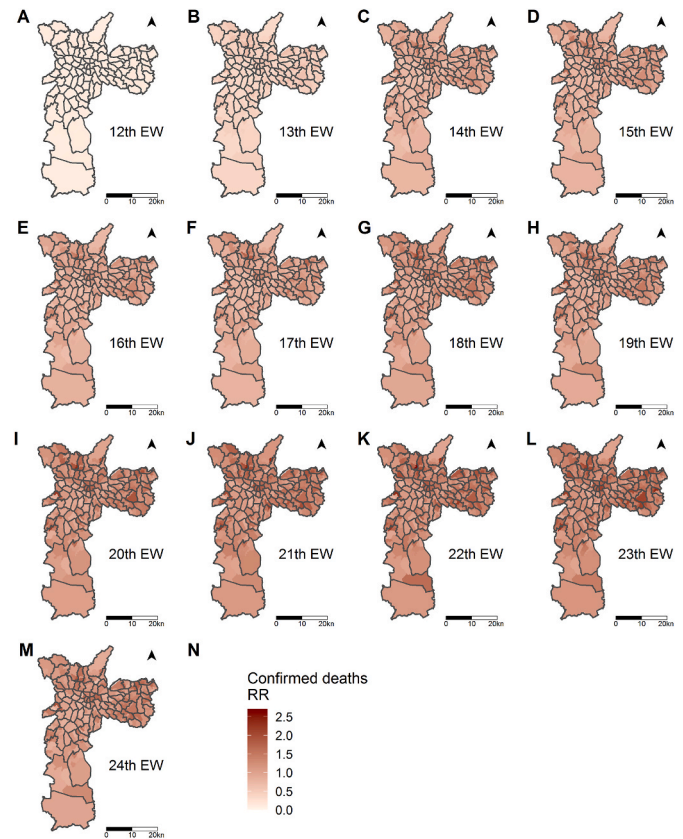
Number and mortality rates (per 100,000 inhabitants in fourteen weeks) of suspected (severe acute respiratory syndrome - U04.9), confirmed (coronavirus infection disease - B34.2), and total (U04.9 + B34.2) COVID-19 deaths, according to e-SIC (São Paulo Electronic Information System) and Tabnet databases, sex, and age. City of São Paulo, 11th to 24th epidemiological weeks, 2020.

	e-SIC database (provided on 6/18/ 2020)		Tabnet database (updated on 7/23/ 2020)		Ratio: Tabnet database/e- SIC database
	N° of deaths	Mortality rate (per 100,000 inhab.)	N° of deaths	Mortality rate (per 100,000 inhab.)	
<b>Male sex</b>					
<b>Confirmed</b>	3350	59.3	3861	68.3	1.2
<b>Suspect</b>	2507	44.4	2228	39.4	0.9
<b>TOTAL</b>	5857	103.6	6089	107.7	1.0
<b>Female sex</b>					
<b>Confirmed</b>	2524	40.6	2906	46.7	1.2
<b>Suspect</b>	2309	37.1	2099	33.8	0.9
<b>TOTAL</b>	4833	77.7	5005	80.5	1.0
<b>Total</b>					
<b>Confirmed</b>	5875	49.5	6768	57.0	1.2
<b>Suspect</b>	4818	40.6	4330	36.5	0.9
<b>TOTAL</b>	10,693	90.1	11,098	93.5	1.0
<b>Total: 0 to 19 years old</b>					
<b>Confirmed</b>	17	0.6	20	0.7	1.2
<b>Suspect</b>	41	1.4	41	1.4	1.0
<b>TOTAL</b>	58	1.9	61	2.0	1.1
<b>Total: 20 to 39 years old</b>					
<b>Confirmed</b>	236	6.2	284	7.5	1.2
<b>Suspect</b>	208	5.5	172	4.5	0.8
<b>TOTAL</b>	444	11.7	456	12.0	1.0
<b>Total: 40 to 59 years old</b>					
<b>Confirmed</b>	1145	35.5	1324	41.1	1.2
<b>Suspect</b>	825	25.6	706	21.9	0.9
<b>TOTAL</b>	1970	61.2	2030	63.0	1.0
<b>Total: 60 years old or older</b>					
<b>Confirmed</b>	4475	241.5	5140	277.3	1.1
<b>Suspect</b>	3740	201.8	3405	183.7	0.9
<b>TOTAL</b>	8215	443.3	8545	461.1	1.0

Data source: Deaths: Mortality Information System of Mortality Information Improvement Program, from the Epidemiology and Information Coordination of the São Paulo Health Secretariat (SMS-SP). (SIM/PRO-AIM/CEInfo/SMS-SP). Population: Fundação Sistema Estadual de Análise de Dados Estatísticos (SEADE). Tabnet database was updated on 7/23/2020 and e-SIC database was provided on August 6, 2020.



**Fig. 3.** Posterior means of the temporal relative risks (RR) of COVID 19. City of São Paulo, 12th to 24th Epidemiological Week, 2020.



**Fig. 4.** Posterior means of the spatiotemporal relative risks (RR) for confirmed COVID-19 deaths. Sample areas of the city of São Paulo, 12th to 24th Epidemiological Week, 2020.

that the degree of underreporting of COVID-19 cases would increase with a decrease in socio-economic status. Therefore, our results confirmed the association between COVID-19 and human development, pointing out the importance of geographic screening in locations with a potential for transmission of local infectious diseases, as a fundamental aspect to coordinate better actions against the pandemic [35].

The low socio-economic condition levels demonstrate not only the vulnerability of the population but also the difficulties health services face with respect to the diagnosis and treatment of the COVID-19. This is similar to the overall fragility expected from health services in Brazil [36] and Latin American countries facing the pandemic [37]. Living conditions may also be strongly influenced by low income in different ways, such as residence in poorer neighbourhoods and poorer housing conditions (particularly confined or overcrowded housing) [38]. The latter condition has been demonstrated to be associated with a greater risk of contagion from several other pathogens such as *Helicobacter pylori* [39], tuberculosis [40], or Epstein-Barr virus [41]. Regarding COVID-19, studies have shown that occupation is an explicit determinant of contagion and a secondary determinant of COVID-19 severity and deaths due to the association between occupational social class and comorbidities [38]. For example, workers such as cleaners, retail staff, teachers, or healthcare workers have a direct impact on COVID-19 incidence [42]. Individuals with underprivileged socio-economic conditions are more prone to be exposed to job stress, including burnout syndrome and unemployment, which may contribute to disrupted immune and inflammatory system responses [43,44] as well as a higher risk of comorbidities with COVID-19 [45]. Until now, both debilitated immunity and the existence of comorbidities are recognised risk factors for COVID-19 severity [38].

We showed that the first cases of deaths occurred in the neighbourhoods with the best socio-economic conditions in the city of São Paulo.

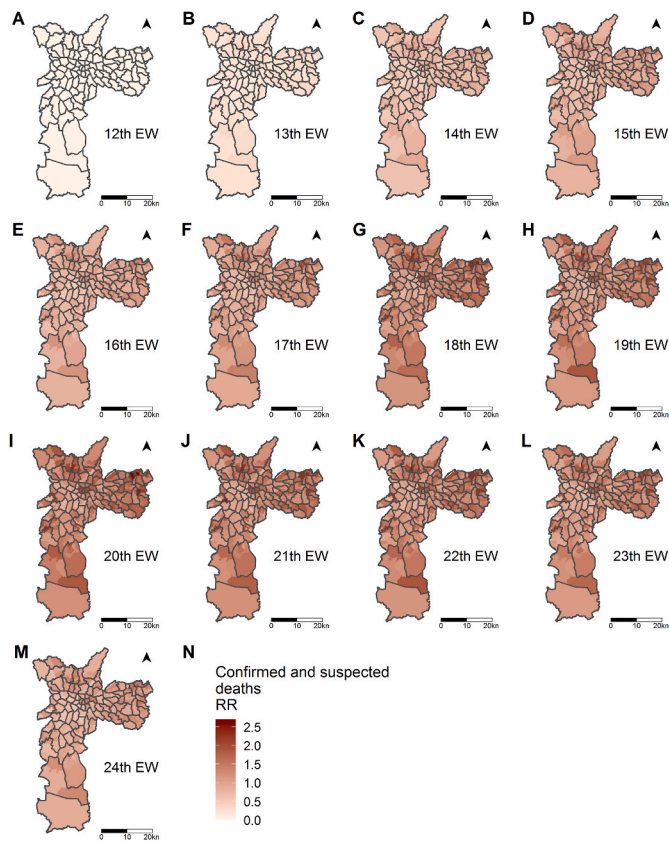


Fig. 5. Posterior means of the spatiotemporal relative risks (RR) for total COVID-19 deaths by sample areas of the city of São Paulo, 12th to 24th Epidemiological Week, 2020.

Table 2

Posterior means of the relative risks (RR) and 95% credibility intervals for the socio-economic covariate (GeoSES) obtained with the spatiotemporal models for confirmed and total COVID-19 deaths. City of São Paulo, 12th to 24th Epidemiology Weeks, 2020.

COVID-19 deaths	Covariate	RR posterior means	95% Credible Interval (CI)	
			0.025 quantil	0.975 quantil
Confirmed	Intercept	0.82	0.78	0.85
	GeoSES	0.75	0.69	0.82
Total	Intercept	0.74	0.71	0.77
	GeoSES	0.67	0.62	0.72

This may be because all the infected subjects had been abroad [46]. In the first two weeks, the best socio-economic conditions was presented as a risk factor. Then, there was a change in the spatial pattern from the fourth week onwards, when the worst socio-economic level became a risk factor. Similarly, Souza et al. [3] showed a higher risk of diagnosed COVID-19 cases in census tracts with higher per capita income in the São Paulo metropolitan region during the early phase of the COVID-19 epidemic. After these first cases in the richest areas, the virus started to circulate in the suburbs of the city, with high population density and worsened sanitary conditions [47]. This probably explains its fast transmission. The city of São Paulo is particularly vulnerable because it is the most populated city in the country, with approximately 12 million inhabitants [48], and it is highly connected with other countries worldwide via its main airport, the São Paulo-Guarulhos International Airport, which is the largest in Brazil, with non-stop passenger flights to 103 destinations across 30 countries [37].

In this study, we used SIM instead of the SIVEP Gripe database,

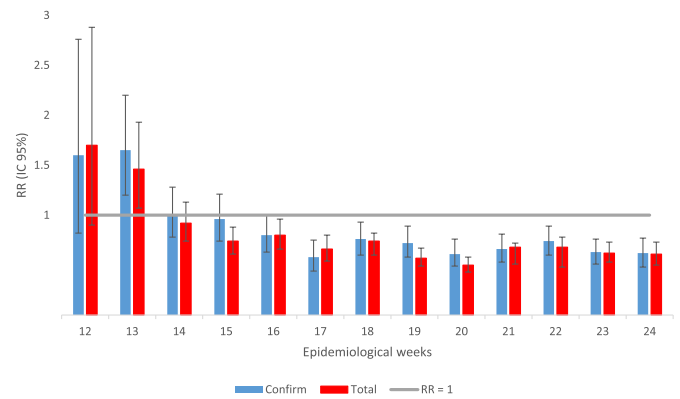


Fig. 6. Posterior means of the relative risks and 95% credible interval for the socio-economic covariate obtained with spatial models for confirmed and total COVID-19 deaths, according to each one of the epidemiologic weeks. City of São Paulo, 12th to 24th Epidemiology Week, 2020.

unlike other studies [3]. The recommendation of the State Health Department [49] to register the notification of death and the monitoring of mortality using SIM, in practice, leads to a time-lapse between the event and the use of information. COVID-19 mortality data were improved in the city of São Paulo with PRO-AIM, using the Laboratory Environment Manager System (GAL) and Flu Surveillance Information System (Sivep Gripe). These characteristics, combined with the possibility of assessing home or unattended deaths, motivated the option of using SIM data (which considered the confirmed and suspected diagnoses in analysing mortality due to COVID-19). It is necessary to emphasise that the recommendation to register confirmed and suspected deaths due to COVID-19 [49] (that motivated the change of the codes assigned by the health services since the beginning of the crisis, and which is applied by SMS- SP to monitor the evolution of events) has not been applied yet.

Our study findings must be considered in the context of several assumptions and data limitations. We associated patients' addresses or postcodes to the area-based socio-economic conditions using geolocalisation. This may provide some insight into the likelihood of exposure to health factors and COVID-19 risks. This approach is frequently used as representative of individual socio-economic conditions. Nevertheless, they are not a perfect picture of the individuals' conditions, and they could underestimate the magnitude of social disproportion related to individual social measures [50]. Rather, they are best employed along with individual-level variables to reflect geographical or aggregate-level risks [38]. We highlight that our spatial analysis is subject to methodological limitations caused by ecological fallacy and the modifiable areal unit problem. These constraints are intrinsic to any spatial analysis that uses aggregated data [51]. Despite these, our study still contributes to healthcare planning measures and to future precision studies focusing on the effects of social health factors on COVID-19 deaths. In addition, one of the strengths of our study was that it dealt with COVID-19 deaths instead of the cases (owing to the better accuracy and coverage of the data). When we consider only the cases, many asymptomatic cases may not be reported and this could hamper the conclusions.

### 5. Conclusions

We used models with spatial and spatiotemporal architectures to investigate the patterns of confirmed and total (confirmed and suspected) COVID-19 deaths in the city of São Paulo. The obtained results, after considering both categories, showed differences regarding the magnitude of the rates and RR. However, there were no differences with respect to the derived conclusions. The risk of mortality due to COVID-19 was the highest between the 20th and 23rd EW, followed by an

apparent stabilisation of the temporal trend. However, we did not rule out a possible future rise in mortality. A high socio-economic level was shown to protect against the risk of mortality due to COVID-19 throughout the study period. However, this was not a uniform pattern, since we identified a shift in the risk of mortality due to COVID-19 in the city of São Paulo over time. We had the highest risk in the best socio-economic contexts during the first two EWs that then shifted to the worst contexts from the 16th EW. Concerning sex and age, men and elderly individuals were at the highest risk of mortality due to COVID-19. Our study corroborated the relationship between COVID-19 mortality and socio-economic conditions, revealing the importance of integrating geographic screening, in areas with higher risk of death, when planning better actions to face the pandemic.

## Funding source

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico [grant numbers 301550/2017-4 to LVB and 306025/2019-1 to FCN]; and the São Paulo Research Foundation (FAPESP) [grant number 2017/10297-1 to CL].

## CRediT authorship contribution statement

**Patricia Marques Moralejo Bermudi:** Data curation, Formal analysis, Methodology, Software, Writing - original draft, Writing - review & editing. **Camila Lorenz:** Investigation, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Breno Souza de Aguiar:** Conceptualization, Data curation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Marcelo Antunes Failla:** Conceptualization, Data curation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Ligia Vizeu Barrozo:** Conceptualization, Investigation, Project administration, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **Francisco Chiaravalloti-Neto:** Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

## Declaration of competing interest

I declare no competing interests.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.tmaid.2020.101945>.

## References

- Maroko AR, Nash D, Pavilonis BT. COVID-19 and inequity: a comparative spatial analysis of New York city and Chicago hot spots. *J Urban Health* 2020. <https://doi.org/10.1007/s11524-020-00468-0>.
- Wang Z, Tang K. Combating COVID-19: health equity matters. *Nat Med* 2020;26. <https://doi.org/10.1038/s41591-020-0823-6>. 458–458.
- Souza WM, Buss LF, Candido D da S, Carrera J-P, Li S, Zarebski AE, et al. Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. *Nat Hum Behav* 2020. <https://doi.org/10.1038/s41562-020-0928-4>.
- Brasil. Ministério da Saúde. Painel coronavírus. 2020. <https://covid.saude.gov.br/>.
- The World Bank DRG. World development indicators. Washington, D.C. 2020.
- Santos NR dos. SUS 30 anos: o início, a caminhada e o rumo. *Ciência Saúde Coletiva* 2018;23:1729–36. <https://doi.org/10.1590/1413-81232018236.06092018>.
- Yang HM, Lombardi Junior LP, Castro FFM, Yang AC. Mathematical model describing CoVid-19 in São Paulo, Brazil – evaluating isolation as control mechanism and forecasting epidemiological scenarios of release. *Epidemiol Infect* 2020;148:e155. <https://doi.org/10.1017/S0950268820001600>.
- Kamel Boulos MN, Geraghty EM. Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbr. *Int J Health Geogr* 2020;19:8. <https://doi.org/10.1186/s12942-020-00202-8>.
- Early Alert. Coronavirus 2020 (dashboard). <https://earlyalert.maps.arcgis.com/apps/opsdashboard/index.html#/20bfb89c8e74c0494c90b1ae0fa7b78>. [Accessed 24 July 2020].
- Johns Hopkins University C. Coronavirus COVID-19 global cases (dashboard). 2020. <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e%0A9ecf6>. [Accessed 24 July 2020].
- World Health Organization. Novel coronavirus (COVID-19) situation (dashboard). [http://healthycbermap.org/WHO\\_COVID19/](http://healthycbermap.org/WHO_COVID19/). [Accessed 24 July 2020].
- Fundação SEADE. População projetada. 2020. <https://produtos.seade.gov.br/produtos/projpop>.
- Instituto Brasileiro de Geografia e Estatística. Metodologia do censo demográfico 2010. Rio de Janeiro, RJ. 2016.
- Brasil. Ministério da Saúde. SINAN – calendário Epidemiológico 2020. <https://portalsinan.saude.gov.br/calendario-epidemiologico-2020>. [Accessed 22 July 2020].
- São Paulo. Prefeitura. Pedidos de informação protocolados à Prefeitura via sistema e-SIC. <http://dados.prefeitura.sp.gov.br/dataset/pedidos-de-informacao-protocolados-a-prefeitura-via-e-sic1>; 2020.
- Guerriero ICZ. Resolução nº 510 de 7 de abril de 2016 que trata das especificidades éticas das pesquisas nas ciências humanas e sociais e de outras que utilizam metodologias próprias dessas áreas. *Ciência Saúde Coletiva* 2016;21:2619–29. <https://doi.org/10.1590/1413-81232015218.17212016>.
- Barrozo LV, Fornaciali M, André CDS, Morais GAZ, Mansur G, Cabral-Miranda W, et al. GeoSES: a socioeconomic index for health and social research in Brazil. *PLoS One* 2020;15:e0232074. <https://doi.org/10.1371/journal.pone.0232074>.
- Duncan GJ, Daly MC, McDonough P, Williams DR. Optimal indicators of socioeconomic status for health research. *Am J Publ Health* 2002;92:1151–7.
- Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Publ Health* 1997;18:341–78.
- Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-INLA. *Spat Spatiotemporal Epidemiol* 2013;4:33–49. <https://doi.org/10.1016/j.sste.2012.12.001>.
- Blangiardo M, Cameletti M. Spatial and spatio-temporal bayesian models with R-INLA. Chichester, UK: John Wiley & Sons, Ltd; 2015. <https://doi.org/10.1002/9781118950203>.
- Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: a principled, practical approach to constructing priors. *Stat Sci* 2017;32:1–28. <https://doi.org/10.1214/16-ST576>.
- Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J Roy Stat Soc B* 2009;71:319–92. <https://doi.org/10.1111/j.1467-9868>.
- R Core Team. R: a language and environment for statistical computing. 2019.
- Governo de São Paulo. Isolation measures in the state of sao Paulo. 2020. <https://www.saopaulo.sp.gov.br/coronavirus/isolamento/>.
- Nussbaumer-Streit B, Mayr V, Dobrescu AI, Chapman A, Persad E, Klerings I, et al. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database Syst Rev* 2020. <https://doi.org/10.1002/14651858.CD013574>.
- CDC. The novel coronavirus pneumonia emergency response Epidemiology team. Vital surveillances: the epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19). <http://weekly.chinacdc.cn/en/>; 2020.
- Pawlowski B, Atwal R, Dunbar RIM. Sex differences in everyday risk-taking behavior in humans. *Evol Psychol* 2008;6. <https://doi.org/10.1177/147470490800600104>. 1474704908006000.
- Raisi-Estabragh Z, McCracken C, Ardissino M, Bethell MS, Cooper J, Cooper C, et al. Non-white ethnicity, male sex, and higher body mass index, but not medications acting on the renin-angiotensin system are associated with coronavirus disease 2019 (COVID-19) hospitalisation: review of the first 669 cases from the UK biobank. *MedRxiv* 2020;2020. <https://doi.org/10.1101/2020.05.10.20096925>. 05.10.20096925.
- Sharma G, Volgman AS, Michos ED. Sex differences in mortality from COVID-19 pandemic. *JACC Case Rep* 2020;2:1407–10. <https://doi.org/10.1016/j.jaccas.2020.04.027>.
- Global Health 5050. COVID-19 sex-disaggregated data tracker. 2020. <http://globalhealth5050.org/covid19>.
- Klein SL, Flanagan KL. Sex differences in immune responses. *Nat Rev Immunol* 2016;16:626–38. <https://doi.org/10.1038/nri.2016.90>.
- Johnson HD, Sholcosky D, Gabello K, Ragni R, Ogonosky N. Sex differences in public restroom handwashing behavior associated with visual behavior prompts. *Percept Mot Skills* 2003;97:805–10. <https://doi.org/10.2466/pms.2003.97.3.805>.
- Bertakis KD, Azari R, Helms LJ, Callahan EJ, Robbins JA. Gender differences in the utilization of health care services - PubMed. *J Fam Pract* 2000;147–52. <https://pubmed.ncbi.nlm.nih.gov/10718692/>. [Accessed 3 August 2020].
- Maciel JAC, Castro-Silva II, Farias MR de. Análise inicial da correlação espacial entre a incidência de COVID-19 e o desenvolvimento humano nos municípios do estado do Ceará no Brasil. *Rev Bras Epidemiol* 2020;23. <https://doi.org/10.1590/1980-5497202000057>.
- Ribas RM, Campos PA de, Brito CS de, Gontijo-Filho PP. Coronavirus Disease 2019 (COVID-19) and healthcare-associated infections: emerging and future challenges for public health in Brazil. *Trav Med Infect Dis* 2020;101675. <https://doi.org/10.1016/j.tmaid.2020.101675>.
- Rodriguez-Morales AJ, Gallego V, Escalera-Antezana JP, Méndez CA, Zambrano LI, Franco-Paredes C, et al. COVID-19 in Latin America: the implications of the first confirmed case in Brazil. *Trav Med Infect Dis* 2020;35:101613. <https://doi.org/10.1016/j.tmaid.2020.101613>.

- [38] Khalatbari-Soltani S, Cumming RG, Delpierre C, Kelly-Irving M. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *J Epidemiol Community Health* 2020. <https://doi.org/10.1136/jech-2020-214297>. jech-2020-214297.
- [39] Webb PM, Knight T, Greaves S, Wilson A, Newell DG, Elder J, et al. Relation between infection with *Helicobacter pylori* and living conditions in childhood: evidence for person to person transmission in early life. *BMJ* 1994;308:750–3. <https://doi.org/10.1136/bmj.308.6931.750>.
- [40] Gupta D, Das K, Balamughesh T, Aggarwal N, Jindal S. Role of socio-economic factors in tuberculosis prevalence. *Indian J Tubercul* 2004;51:27–31.
- [41] Gares V, Panico L, Castagne R, Delpierre C, Kelly-Irving M. The role of the early social environment on Epstein Barr virus infection: a prospective observational design using the Millennium Cohort Study. *Epidemiol Infect* 2017;145:3405–12. <https://doi.org/10.1017/S0950268817002515>.
- [42] Koh D. Occupational risks for COVID-19 infection. *Occup Med (Chic Ill)* 2020;70:3–5. <https://doi.org/10.1093/occmed/kqaa036>.
- [43] Berger E, Castagné R, Chadeau-Hyam M, Bochud M, D'Errico A, Gandini M, et al. Multi-cohort study identifies social determinants of systemic inflammation over the life course. *Nat Commun* 2019;10:773. <https://doi.org/10.1038/s41467-019-08732-x>.
- [44] Nakata A. Psychosocial job stress and immunity: a systematic review. 2012. p. 39–75. [https://doi.org/10.1007/978-1-62703-071-7\\_3](https://doi.org/10.1007/978-1-62703-071-7_3).
- [45] Kivimäki M, Kawachi I. Work stress as a risk factor for cardiovascular disease. *Curr Cardiol Rep* 2015;17:74. <https://doi.org/10.1007/s11886-015-0630-8>.
- [46] da Saúde Ministério. <https://www.saude.gov.br/noticias/agencia-saude/46435-brasil-confirma-primeiro-caso-de-novo-coronavirus>; 2020.
- [47] Silva JB da, Muniz AMV. Pandemia do coronavírus no brasil: impactos no território cearense. *Espaço e Econ*; 2020. <https://doi.org/10.4000/espacoeconomia.10501>.
- [48] IBGE. Cidades e Estados. 2020. <https://www.ibge.gov.br/cidades-e-estados/sp.html>.
- [49] Brasil. Ministério da Saúde. Orientações para codificação das causas de morte no contexto da COVID-19. 2020.
- [50] Lamy S, Molinié F, Daubisse-Marliac L, Cowppli-Bony A, Ayrault-Piault S, Fournier E, et al. Using ecological socioeconomic position (SEP) measures to deal with sample bias introduced by incomplete individual-level measures: inequalities in breast cancer stage at diagnosis as an example. *BMC Publ Health* 2019;19:857. <https://doi.org/10.1186/s12889-019-7220-4>.
- [51] Subramanian SV, Jones K, Kaddour A, Krieger N. Revisiting Robinson: the perils of individualistic and ecologic fallacy. *Int J Epidemiol* 2009;38:342–60. <https://doi.org/10.1093/ije/dyn359>.