# Multi-generational SIR modeling: Determination of parameters, epidemiological forecasting and age-dependent vaccination policies

Eduardo Lima Campos [a, b], Rubens Penha Cysne [a], Alexandre L. Madureira [c, a, *], Gélcio L.Q. Mendes [d]

[a] EPGE Brazilian School of Economics and Finance (FGV EPGE), Rio de Janeiro, RJ, Brazil
[b] ENCE - Escola Nacional de Ciências Estatísticas (ENCE/IBGE), Rio de Janeiro, RJ, Brazil
[c] Laboratório Nacional de Computação Científica, Petrópolis, RJ, Brazil
[d] INCA - Brazilian National Cancer Institute, Coordination of Assistance, Rio de Janeiro, RJ, Brazil

## ARTICLE INFO

## ABSTRACT

We use an age-dependent SIR system of equations to model the evolution of the COVID-19. Parameters that measure the amount of interaction in different locations (home, work, school, other) are approximated from in-sample data using a random optimization scheme, and indicate changes in social distancing along the course of the pandemic. That allows the estimation of the time evolution of classical and age-dependent reproduction numbers. With those parameters we predict the disease dynamics, and compare our results with out-of-sample data from the City of Rio de Janeiro. Finally, we provide a numerical investigation regarding age-based vaccination policies, shedding some light on whether is preferable to vaccinate those at most risk (the elderly) or those who spread the disease the most (the youngest). There is no clear upshot, as the results depend on the age of those immunized, contagious parameters, vaccination schedules and efficiency.

## 1. Introduction

Several months after the onset of the COVID-19 pandemic it became clear that good numerical modeling is useful to understand the dynamics of the disease. In particular, simple compartmental epidemiological models are convenient, depending on few parameters and still capturing with reasonable precision essential aspects of the dynamics of infectious diseases (Brauer, 2017; Hethcote, 2000; Wang et al., 2021).

The Susceptible-Infected-Recovered (SIR) model in its simplest form depends on two parameters, the average number of *sufficient contacts* (those sufficient for transmission) and the *mean waiting period* during which patients are contagious. It relies on the assumption that the population is homogeneous both in terms of behavior and biological susceptibility. It also

---

assumes, in its simplest instance, that demographic aspects as age structure, birth and death rates do not matter for the dynamics of the disease.

Some of the above assumptions appear to hold well in the case of the COVID-19 pandemic. For instance, demographic changes appear to be of less importance due to difference in time scales, compared with the evolution of the pandemic. However, dropping the homogeneity hypothesis allows a more realistic model, as age and location of contacts are important factors in understanding the dynamics of the disease. For instance, the disease's severity and lethality is age-dependent, among other factors (Bhopal & Bhopal, 2020; Jin et al., 2020; Verity et al., 2020; Wu & McGoogan, 2020).

Mathematical epidemiology has a somewhat long history, with (Hamer, 1906; M'Kendrick, 1925; Kermack, McKendrick, & Walker, 1927) being among the earliest contributions. See (Brauer, 2017; Hethcote, 2000) for a thorough review of mathematical modeling of infectious diseases, and (Albani et al., 2021a, 2021b·bib_Albani_et_al_2021a·bib_Albani_et_al_2021b; Iannelli & Milner, 2017; Klepac & Caswell, 2011; Klepac et al., 2009; bib_Li_et_al_2020Li, Yang, & Martcheva, 2020, pp. 1–21; Prem et al., 2020; Sun, 2010; Towers & Feng, 2012) for multi-group models. Some other models consider spatial aspects of infectious diseases, as (Bertaglia & Pareschi, 2020; Besse & Faye, 2020; Lang, De Sterck, Kaiser, & Miller, 2018; Paeng & Lee, 2017; Peixoto, Marcondes, Peixoto, & Oliva, 2020; Takács & Hadjimichael, 2019; Viguerie et al., 2020). Theoretical aspects of models are covered in (Kuniya & Wang, 2018; Wen, Ji, & Li, 2018; Wu & Zou, 2016). For the control related models for COVID-19 see for instance (Perkins & España, 2020) and references therein.

After the onset of the COVID-19 pandemic, a massive amount of articles related to the topic materialized. Several articles considered compartmental models to forecast different scenarios (Albani et al., 2021a, 2021b·bib_Albani_et_al_2021a·bib_Albani_et_al_2021b; Amaku et al., 2021a, 2021b·bib_Amaku_et_al_2021a·bib_Amaku_et_al_2021b; Atkeson, 2020; Coelho et al., 2020; Kimathi, Mwalili, Ojiambo, & Gathungu, 2021; Komatsu & Menezes-Filho, 2020; Prem et al., 2020; Viguerie et al., 2020; Walker et al., 2020).

Socioeconomic costs of imposing social distancing are combined with compartmental models in (Acemoglu, Chernozhukov, Werning, & Whinston, 2020; Alvarez, Argente, & Lippi, 2020; Atkeson, 2020), and the interplay between economics and individual decisions are considered in (Borelli & Góes, 2020; Brotherhood, Kircher, Santos, & Tertilt, 2020; Eichenbaum, Rebelo, & Trabandt, 2020). See (Kremer, 1996; d'Onofrio & Manfredi, 2020, pp. 185–203; Funk et al., 2015; Manfredi and d'Onofrio, 2013; Perrings et al., 2014; Soofi, Najafi, & Karami-Matin, 2020) for a description of behavioral- and economic-epidemiology ideas. Also, macro-economic aspects of the pandemic appears, e.g., in (Guerrieri, Lorenzoni, Straub, & Werning, 2020).

Population heterogeneity and its influence on herd immunity is considered in (Aguas et al., 2020; 2020Britton, Ball, & Trapman, ; Cui, Zhang, & Feng, 2019; Gomes et al., 2020). Also, different aspects of vaccination are considered in the literature (Albani et al., 2021b; Amaku et al., 2021a; d'Onofrio & Manfredi, 2020, pp. 185–203; Meehan et al., 2020; Chen & Toxvaerd, 2014; González & Villena, 2020; Jia et al., 2020; May & Anderson, 1984). Finally, alternative approaches, as agent-based, statistical models and a combination of these techniques with compartmental models (specially to gather data), are also commonly employed (Amaral, Casaca, Oishi, & Cuminato, 2021; Bendtsen Cano et al., 2020; Calvetti, Hoover, Rose, & Somersalo, 2020; Donnat & Holmes, 2004; Ferguson, Laydon, & Nedjati-Gilani, 2020; Roda, Varughese, Han, & Li, 2020).

In the present paper, we consider a modification of the age-structured SIR model (Britton et al., 2020; Towers & Feng, 2012), in particular allowing for age-dependent probabilities of patients being sub-clinical (asymptomatic or symptomatic patients that does not require hospitalization) or clinical (Albani et al., 2021b; Prem et al., 2020). We use contact matrices (Prem, Cook, & Jit, 2017) which differentiates not only age but also location (home, work, school, other); see also (Kimathi et al., 2021). Random searches approximate some key parameters that are related to social distancing in different locations, computing their values for some Brazilian regions. Following the approximation of those parameters, we are able to forecast long term best/worst case scenarios. We gauge the efficiency of the scheme comparing our predictions with available data using in-sample/out-of-sample testing.

We next use the model to compare age-dependent vaccination policies (Meehan et al., 2020). In a simple setting, we compare vaccinating the young or the older population, since the model allows the immunization of specific age groups. The outcome depends on the values of certain parameters, including how aggressive is the transmission and how late after the start of the pandemic the vaccinations take place.

In general, for a short term period, vaccination prioritizing the elderly population is the best option with respect to the total number of deaths. But that might not be the case in the long run if extra vaccine supplies are not provided.

**Disclaimer:** It is not our intention to propose real-life epidemiological policies or to provide predictions upon which policies could be developed. Our intention is solely to investigate possible patterns of the disease evolution and raise possible alternatives in terms of vaccination policies.

## 2. The model

In the spirit of compartmental modeling (Hethcote, 2000; Wang et al., 2021), we divide the whole population among *susceptible* (**S**), *infected* (**I**), *removed* (**R**). Due to the short time scale of the disease, we can assume that demographic changes are not relevant. The number of deaths are estimated *a posteriori*, from the removed **R**. The above quantities are 16-

dimensional vectors, stratified by age, running from 0 to 4 (first components), 5–9 (second components), etc, up to 75-above (16th components). There are $P_i$ individuals for each age group, and the total population $P = \sum_{i=1}^{16} P_i$ is constant.

The equations governing the dynamics of the disease is as follows,[1] for $i = 1 \ldots, 16$ and $t \in \mathbb{N}$:

$$
\begin{aligned}
S_i(t+1) &= S_i(t) - \beta_i(\boldsymbol{I})S_i(t), \\
I_i(t+1) &= I_i(t) + \beta_i(\boldsymbol{I})S_i(t) - \gamma I_i(t), \\
R_i(t+1) &= R_i(t) + \gamma I_i(t),
\end{aligned}
\tag{1}
$$

plus initial conditions. The model is "conservative" in the sense that $S_i + I_i + R_i = P_i$ is constant for all ages $i$.

The term $\beta_i S_i$ is the rate of new infections of the i-population, and we remark that $\beta$ depends on the number of patients of all ages, making the model coupled and nonlinear. These patients are removed from their condition at rate $\gamma I_i$ (they either recover or die). We further divide the infectious into two groups, those that are asymptomatic or *sub-clinical* $I_i^{sc} = (1 - \rho_i)I_i$ in the sense that they do not require medical care. The *clinical* group $I_i^c = \rho_i I_i$ corresponds to those that get ill, requiring medical attention. The age-dependent fractions of those who become sub-clinical or clinical are $\rho_i \in [0, 1]$.

To fully describe the model, we ought to define $\beta$ and the other parameters.

(i) the components of the incidence function $\beta$ are given by the formula

$$
\beta_i(\boldsymbol{I}) = \sum_{j=1}^{16} \frac{\alpha^{sc} I_j^{sc} + \alpha^c I_j^c}{P_j} C_{i,j}^e = \sum_{j=1}^{16} \left(\alpha^{sc}(1 - \rho_j) + \alpha^c \rho_j\right) \frac{I_j}{P_j} C_{i,j}^e,
$$

where $C^e$ is an *effective contact matrix* that measures the number of contacts between age groups. It depends on the *amount* and *choices* of lock-down imposed, as we describe further ahead. The age-independent parameters $\alpha^{sc}$ and $\alpha^c$ correspond to the fraction of those that are infective, having the potential to infect others due to social behavior or biology. For convenience, we also write the vector $\beta(\boldsymbol{I}) = C^e D \boldsymbol{I}$, where $D$ is a diagonal matrix defined by

$$
D_{jj} = \frac{\alpha^{sc}(1 - \rho_j) + \alpha^c \rho_j}{P_j}.
$$

(ii) The time dependent effective contact matrix $C^e$ is defined by

$$
C^e(t) = \beta_h(t)C^{\text{home}} + \beta_w(t)C^{\text{work}} + \beta_s(t)C^{\text{school}} + \beta_o(t)C^{\text{other}},
\tag{2}
$$

where the exogenous data $C^{\text{home}}$, $C^{\text{work}}$, $C^{\text{school}}$, $C^{\text{other}}$ are $16 \times 16$ matrices, where $C_{ij}^\ell$ indicates the average number of *different people* from the age group $j$ that someone from the age group $i$ contacted per day at $\ell \in \{\text{home, work, school, other}\}$, as compiled by (Mossong et al., 2008; Prem et al., 2017). The parameters $\beta_\ell$ control the fraction of contacts that are "adequate", in the sense that they lead to infections.

(iii) Each component of $\rho$ gives the risk that an infectious will become a sub-clinical or a clinical case. Since COVID-19 symptoms are more aggressive for older patients, $\rho_i$ grows with $i$.

(iv) $\gamma = 1 - e^{-1/d_I}$ is the daily probability that someone stays infected, where $d_I$ is the average duration of infection (Prem et al., 2017). For simplicity, we assume that the period in which one is infective is independent of age and disease severity. These hypotheses are commonly, but not always, assumed in the literature (Acemoglu et al., 2020; Albani et al., 2021a, 2021b'bib_Albani_et_al_2021a'bib_Albani_et_al_2021b; Britton et al., 2020; Prem et al., 2020); see (Levin et al., 2020; Walsh et al., 2020; Zhou, She, Wang, & Ma, 2020) for related discussions based on medical data.

(v) The contagious parameters $\alpha^{sc}$ and $\alpha^c$ correspond to fractions of sub-clinical and clinical infected patients that still spread the virus.

### 2.1. Reproduction and replacement numbers

The *basic reproduction number* $\mathcal{R}_0$ is understood as *the secondary cases produced by an infective patient introduced in a totally susceptible population* (1002Barril, Calsina, Cuadrado, & Ripoll, ; Delamater, Street, Leslie, Yang, & Jacobsen, 2019;

---

[1] Note that we are *not* using the Einstein's summation convention of summing up repeated indices.

Diekmann, Heesterbeek, & Metz, 1990; Hethcote, 2000). Hence, suppose that the whole population with distribution $\boldsymbol{P}$ is susceptible ($\boldsymbol{S} = \boldsymbol{P}$), and we represent the infective patients by $\hat{\boldsymbol{I}}$, a vector belonging to the canonical basis $\{\boldsymbol{e}_1, ..., \boldsymbol{e}_{16}\}$. The number of patients in the first day belonging to the $i$-th group is

$$\sum_{k=1}^{16} \delta_{ik}\beta_k(\hat{\boldsymbol{I}})P_k = \sum_{k,j,\ell=1}^{16} \delta_{ik}P_k C_{kj}^e D_{jl}\hat{\boldsymbol{I}}_l = \sum_{m,j,\ell=1}^{16} \mathbb{P}_{im}C_{mj}^e D_{jl}\hat{\boldsymbol{I}}_l$$

where $\delta_{ik}$ denotes the Kronecker delta, which equals one if $i = j$ and zero otherwise, and we define $\mathbb{P}$ as the diagonal matrix such that $\mathbb{P}_{kk} = P_k$. Let $A = \mathbb{P}C^e D$. Then the vector containing the contaminated individuals in the first day is $A\hat{\boldsymbol{I}}$, and we can bound its Euclidean norm as

$$\|A\hat{\boldsymbol{I}}\|_2 \leq \lambda_{\max}\|\hat{\boldsymbol{I}}\|_2 = \lambda_{\max}, \tag{3}$$

where $\lambda_{\max}$ is the largest eigenvalue of $A$. Note that only the fraction $(1 - \gamma)^{d-1}\hat{\boldsymbol{I}}$ of the infective outsider remains infected on the d-th day. This fraction infect less than $(1 - \gamma)^{d-1}\lambda_{\max}$, as in (3), and adding up all days, we gather that the total number of infective agents is bounded by the *basic reproduction number*

$$\mathcal{R}_0 := \sum_{d=1}^{\infty} (1 - \gamma)^{d-1}\lambda_{\max} = \frac{1}{\gamma}\lambda_{\max}. \tag{4}$$

Reasoning in a similar fashion, we define the *replacement number*

$$\mathcal{R}_t := \frac{1}{\gamma}\lambda_{\max}(t), \tag{5}$$

where $\lambda_{\max}(t)$ is defined as the maximum eigenvalue of the matrix $\mathbb{S}C^e D$ and $\mathbb{S}$ is the diagonal matrix with the diagonal given by $S_j$ for $j = 1, ..., 16$.

The above definitions of $\mathcal{R}_0$ and $\mathcal{R}_t$ rely on the Euclidean and spectral norms. However, if one is interested in considering age-dependent quarantine or vaccination policies, it might be interesting to consider a slightly different approach. An infective patient of the age group $\ell \in \{1, ..., 16\}$ would contaminate, on a single day,

$$\beta(\boldsymbol{e}_\ell) \cdot \boldsymbol{P} = C^e D\boldsymbol{e}_\ell \cdot \boldsymbol{P} = \sum_{i=1}^{16} C_{i\ell}^e D_{\ell\ell}P_i. \tag{6}$$

As before, adding up the days and considering that at the jth day there will be a fraction $(1 - \gamma)^{d-1}\boldsymbol{e}_\ell$ of the infective outsider, we gather that the total number of infective agents is the *age-dependent basic reproduction number*

$$\mathcal{R}_{0,\ell} := \sum_{d=1}^{\infty} (1 - \gamma)^{d-1}(\boldsymbol{P}^T C^e D)_\ell = \frac{1}{\gamma}(\boldsymbol{P}^T C^e D)_\ell. \tag{7}$$

Similarly, we define the corresponding *age-dependent replacement number*

$$\mathcal{R}_{t,\ell} := \frac{1}{\gamma}(\boldsymbol{S}^T C^e D)_\ell. \tag{8}$$

The above number $\mathcal{R}_{t,\ell}$ indicates how many people would be contaminated by an infective patient from the age group $\ell$ introduced in the community at time $t$.

## 2.2. Modeling the lethality

Mortality is not part of the dynamics,[2] resulting from the number of infectious. We postulate that only those clinically infected die, at an age-dependent rate. So, among the "recovered" patients,

$$d(t) = \gamma\mu(t)\boldsymbol{w}_d \cdot \boldsymbol{I}^c(t) \tag{9}$$

die, where the constant vector $\boldsymbol{w}_d$ is the age-dependent lethality weights of the disease (Levin et al., 2020), and the *lethality strength* $\mu$ is a scalar that captures how lethality varies with time. Note that the lethality *proportion* between different ages is constant.

---

[2] We assume that the recovered patients do not get reinfected.

### 2.3. Data

The transmission parameters $\alpha^{sc}$ and $\alpha^c$ are such that the transmission rate of symptomatic is higher than those asymptomatic (Byambasuren et al., 2020; Prem et al., 2020), and we stipulate that the proportion of symptomatic patients follows the same pattern of the hospitalized patients in (Verity et al., 2020). However, the proportion of asymptomatic patients is far from being settled (Byambasuren et al., 2020; Heneghan, Brassey, & Jefferson, 2020), with estimates ranging from 5% to 80%. Also, we assume that profile of the lethality rates follow Fig. 1, based on (Brazeau et al., 2020; Verity et al., 2020).

The contact matrices $C^{home}$, $C^{work}$, $C^{school}$, $C^{other}$ are from (Prem et al., 2017), and population data, as seen in Fig. 10, is from the Instituto Brasileiro de Geografia e Estatística (IBGE). The average number of days $d_I$ that a patient is infective is set to 12.

We use the available data (number of new cases and deaths) from each location to estimate the lethality strength $\mu(t)$, and $\beta_h(t)$, $\beta_w(t)$, $\beta_s(t)$, $\beta_o(t)$, which are necessary to compute the effective matrix $C^e$. These functions are transient since not only individuals change behavior with time (Brotherhood et al., 2020), but also lethality changes with medical conditions and virus mutations.

For a simple SIR model, the available data uniquely determines the parameters (Bakhta, Boiveau, Maday, & Mula, 2020). In general, that is not the case for our model, for instance if one of the four contact matrices vanish or if $C^{work} = C^{other}$. We reduce the possibility of parameter indeterminations by assuming that $\beta_w = \beta_o$, i.e., the behavior at work is equal to the behavior at other locations. We also make then some further simplifying hypotheses. First, we assume that $\beta_h$ is constant, as there is no reason to expect that the amount of contacts at home varies significantly with time. The second assumption is that $\beta_s(t)$ is identically zero, since schools were closed during the period of data gathering.

Consider that data is available for the period $\{1, ..., T\} \subseteq \mathbb{N}$, and let and $\mathcal{N}_{\mathcal{I}}^{data}(j)$ and $\mathcal{N}_{\mathcal{D}}^{data}(j)$ be the number of new cases and new deaths at the $j$-th day. For simplicity, assume that $T = N\delta t$ where $\delta t = 10$, and consider the partition $\mathcal{T}_{\delta t}$ of $\{1, ..., T\}$ $= \cup_{j=1}^{N} I_j$ where

$$\mathcal{T}_{\delta t} = \{I_j : j = 1, ..., N\}, \qquad I_j = \{(j-1)\delta t + 1, ..., j\delta t\}$$

Let $P_0(\mathcal{T}_{\delta t})$ be the space of piecewise constant functions with respect to the above partition. To determine $\beta_w$ we pose the problem of finding $(\beta_h, \beta_w)$ that solves

$$\min_{(\hat{\beta}_h, \hat{\beta}_w) \in \mathbb{R} \times P_0(\mathcal{T}_{\delta t})} J(\hat{\beta}_h, \hat{\beta}_w), \qquad \text{where} \qquad J(\hat{\beta}_h, \hat{\beta}_w) = \frac{\left\| \mathcal{N}_{\mathcal{I}}^{sir}(\hat{\beta}_h, \hat{\beta}_w) - \mathcal{N}_{\mathcal{I}}^{data} \right\|_{\mathbb{R}^T}}{\left\| \mathcal{N}_{\mathcal{I}}^{data} \right\|_{\mathbb{R}^T}}, \tag{10}$$

and $\| \cdot \|_{\mathbb{R}^T}$ is the Euclidean norm in $\mathbb{R}^T$. Above, $\mathcal{N}_{\mathcal{I}}^{sir}(\hat{\beta}_h, \hat{\beta}_w)(t)$ is the number of new infected clinical patients that the SIR model (1) yields at the $t$-th day if one replaces $\beta_h$ and $\beta_w$ by $\hat{\beta}_h$ and $\hat{\beta}_w$ in (2).

The above problem has finite dimension $N + 1$ since any function $\beta \in P_0(\mathcal{T}_{\delta t})$ can be written as

$$\beta_w(t) = \sum_{j=1}^{N} b_j \chi_{I_j}(t) \tag{11}$$

for some unknown coefficients $b_1, ..., b_N \in \mathbb{R}$. Here, $\chi_{I_j}(t)$ is the characteristic function of $I_j$, which equals one if $t \in I_j$, and zero otherwise. Note that the coefficients uniquely define a function in $P_0(\mathcal{T}_{\delta t})$ and vice-versa.

To solve (10) we employ a random optimization approach (Matyaš, 1965; Sarma, 1990)
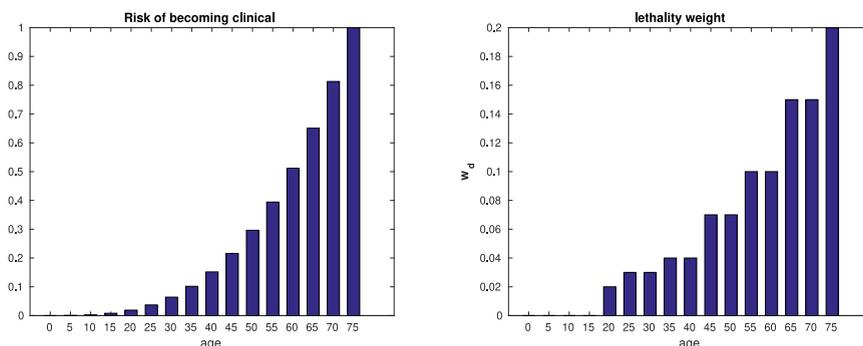
(i) Choose a initial guess $(\beta_h, \beta_w)$



**Fig. 1.** Probability of being clinical among infectives $\rho$ (left) and lethality weight $\mathbf{w}_d$ (right).
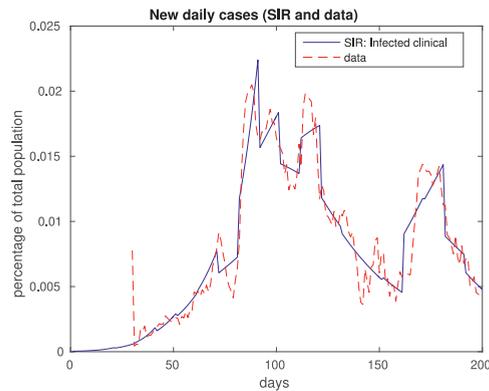
**Fig. 2.** In this figure we plot the 7-day moving average data on new daily cases. The dashed red plot corresponds to real data, and the solid blue plot depicts clinical cases modeled by SIR. Above, and in all figures that follow, what we call data is actually a 7-day moving average of the real data.

(ii) While some convergence criteria is not reached
  (a) Define $(\hat{\beta}_h, \hat{\beta}_w)$ from $(\beta_h, \beta_w)$ by adding noise
  (b) If $J(\hat{\beta}_h, \hat{\beta}_w) < J(\beta_h, \beta_w)$, then $(\beta_h, \beta_w) \leftarrow (\hat{\beta}_h, \hat{\beta}_w)$

The estimation of the lethality strength $\mu$ is based on the number of daily deaths data $\mathcal{N}_{\mathcal{D}}^{\text{data}}$. The total number of deaths simulated by the SIR model is given by $\gamma \mu(t) \boldsymbol{w}_d \cdot \boldsymbol{I}^c(t)$. Equating both quantities we compute

$$\mu(t) = \frac{\mathcal{N}_{\mathcal{D}}^{\text{data}}(t)}{\gamma \boldsymbol{w}_d \cdot \boldsymbol{I}^c(t)}. \tag{12}$$

Of course, parameter determination could be pursued using more sophisticated methods, e.g. (Albani et al., 2021a).

## 3. Simulation of the dynamics of the disease

In what follows we show the results coming out of parameter estimations. We then show results addressing how well the model predicts the number of new cases. We compare with real daily data for new cases and deaths,[3] and, to smooth out oscillations, we present the data using a 7-day moving average.

One of the hurdles for compartmental models is the definition of initial conditions. Indeed, both the exact "first day" of the disease and how many people were infected in that first day are unknown. As a initial condition of our model, we assume that the first infection occurred 30 days before the official recordings. Indeed, the first day of the official data logs 489 cases; cf. (Lourenço et al., 2020), who speculates that the epidemic in Italy and UK started one month before the first reported death. We also stipulate that there were 10 infectious at each age group at day one.

Such uncertainty of initial conditions makes the results unreliable at the initial periods of simulations. In particular, the computation of (12) is not feasible for the lack of data. Thus, we impose $\mu(j) = 1$ at the initial stages.

### 3.1. Parameter estimations, and prediction analysis

We consider results for the City of Rio de Janeiro, henceforward simply called *Rio*. The results presented are for the sum of all ages:

$$S(t) = \sum_{j=1}^{16} S_j(t), \qquad I^{sc}(t) = \sum_{j=1}^{16} I_j^{sc}(t),$$

$$I^c(t) = \sum_{j=1}^{16} I_j^c(t), \qquad R(t) = \sum_{j=1}^{16} R_j(t).$$

We recall the assumption that the contacts might lead to infections at home are time independent, and the contacts at work and other locations are identical, i.e., $\beta_w = \beta_o$. Also, schools are closed, thus $\beta_s = 0$.

---

[3] All data were from taken from https://covid.saude.gov.br/, an official site from the Brazilian's Government Department of Health.
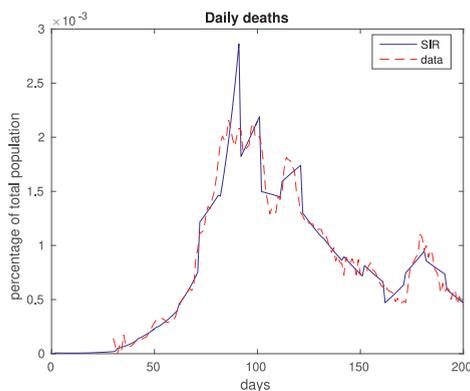
**Fig. 3.** This figure is related to fatalities, displaying daily deaths. Again, real data is in dashed red, and computed daily deaths by the model come in solid blue.
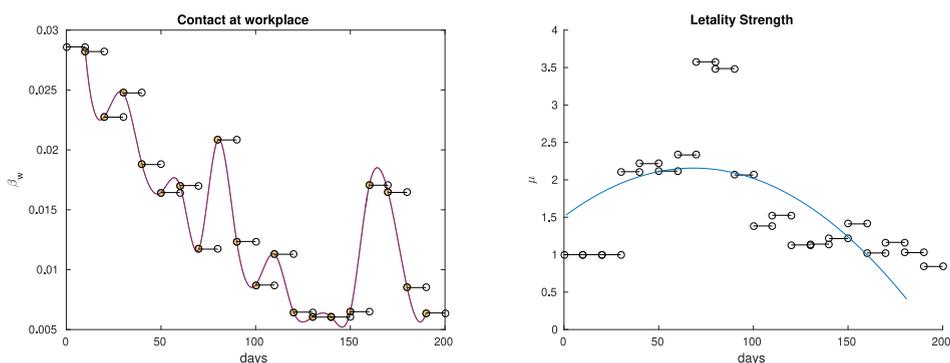


**Fig. 4.** On the left we plot the estimated values of $\beta_w$, except for the first one (that has a value roughly ten times bigger than the others). For visualization purposes we interpolated the discontinuous function $\beta_w$ using cubic splines. On the right we display the values of the lethality strength $\mu$, defined in (12). The blue curve is the least square quadratic curve.

We compare in Figs. 2 and 3 the results of our in-sample approximation of the dynamics of the disease. Note that the actual daily numbers of new cases (Fig. 2) and deaths (Fig. 3) are well approximated by the model.

In Fig. 4 we display the values of $\beta_w$ and $\mu$, estimated by our random search. We see in particular that $\beta_w$ oscillates following the "macroscopic" oscillations of daily cases. Regarding the lethality strength $\mu$, note that it seems to decrease after day 90, in
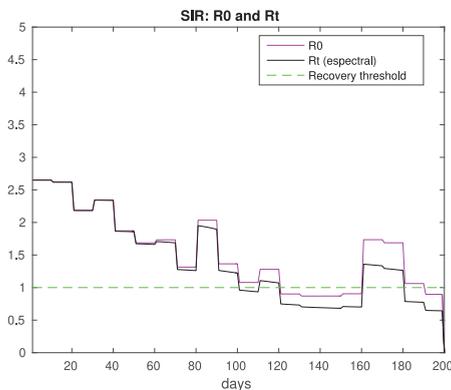


**Fig. 5.** Reproduction and replacement numbers, $\mathcal{R}_0$ and $\mathcal{R}_t$. Note that $\mathcal{R}_0$ is transient since it depends on the time dependent contact matrix. We also draw a line at $y = 1$ highlighting the recovering threshold. Comparing with Fig. 2, we see that there is a correspondence between $\mathcal{R}_t$ crossing the threshold (up or down) and the number of new cases increasing or decreasing.
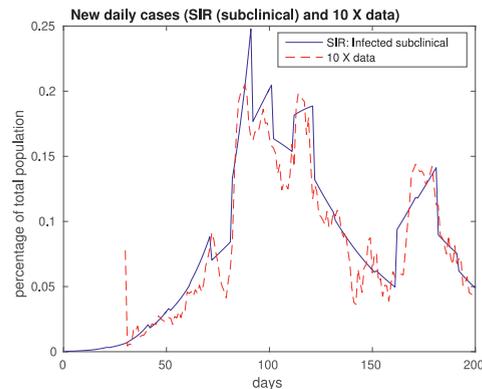
**Fig. 6.** Number of sub-clinical cases in solid blue, corresponding roughly to ten times the number of official cases in dashed red.

line with what was presented by (Dennis, McGovern, Vollmer, & Mateen, 2020; Horwitz et al., 2020). Such aspect is captured by the concave least square parabola.

Two important pieces of information coming out of the model are the values of the reproduction and replacement numbers $\mathcal{R}_0$ and $\mathcal{R}_t$. Their computations follow easily from (4) and (5), and we display the results in Fig. 5. Note that values of $\mathcal{R}_0$ is characterized only by the population's profile and the contact matrix $C^e$. On the other hand, $\mathcal{R}_t$ depends also on the dynamics of the disease through the susceptible population. That is important since even if $\mathcal{R}_0 > 1$, having $\mathcal{R}_t$ smaller than one is enough to have a declining number of infectious.

Another issue that is worth discussing is that we fit the number of clinically infective patients $I^c$ to the data. However, we discussed nothing thus far about the sub-clinical patients $I^{sc}$. It turns out that the number of sub-clinical patients is, in this case, roughly ten times the number of registered patients, as shown in Fig. 6. Such difference resonates well with claims that there is a significant sub-notification. For instance (Havers et al., 2020), reports that the number of cases in the US might be at least 10 times of what is registered. It would not come as a surprise if this difference is even greater in Rio. Also, the WHO states that 80% of infections are mild or asymptomatic,[4] in particular among children, as discussed by (DeBiasi & Delaney, 2770; Han et al., 2019). Recall that children are often not tested.

Next, we investigate numerically if this model can make reasonable predictions. To do so, after the day 200, we run the simulations based on an in-sample (days 1–200) calculated $\beta_w$ and $\mu$, and compare with available data. We perform experiments considering the best and worst scenarios, i.e. lowest and highest number of cases and deaths. We also consider an intermediary case. The main idea shares similarities with some Value at Risk analyses used in risk management, where the past dynamics of investments help in predicting the future; see (Jorion, 2006).

We make the predictions by going back $j$ 10-day periods and performing simulations with the highest and lowest values of $\beta_w$ and $\mu$ in the period $200 - 10j$ and 200. Choosing $j = 4$, we show in Figs. 7–9 the predicted number of daily and accumulated cases and deaths, with 70 days in advance. Note that, for most of the period of interest, the data is in between the best/worst case scenarios. We also plot the intermediary case, corresponding to the choice of $\beta_w$ and $\mu$ being the average of the extreme values, often yielding more accurate predictions.

## 4. Vaccination

We explore in this section simple immunization policies, shedding some light on the following question: if an age-based vaccination strategy should be implemented, who should be protected first? Certainly, real-life policies are never as simple as that, since other aspects as economics, logistics, ethics, politics, etc, play significant roles.[5] Not only that, but also it is not even clear that vaccinations schemes should be devised based on age groups (Pastor-Satorras & Vespignani, 2002). However, vaccination policies based on age are easy to implement in practice.

Based on the age-dependent reproduction and replacement numbers $\mathcal{R}_{0,\ell}$ and $\mathcal{R}_{t,\ell}$, we find out which age group contaminates most people. Consider for instance the dynamics of the disease in Rio, with a population profile as in Fig. 10.

Based on the evolution of the COVID-19 in the first 200 days in Rio, we computed the age-dependent replacement numbers $\mathcal{R}_{t,\ell}$, displaying the results in Fig. 11. Note that as far as spreading the disease is concerned, individuals of younger age are the most dangerous. And individuals above 60, which are the ones most affected by the disease, contribute less for its spreading.

---

[4] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/q-a-similarities-and-differences-covid-19-and-influenza.

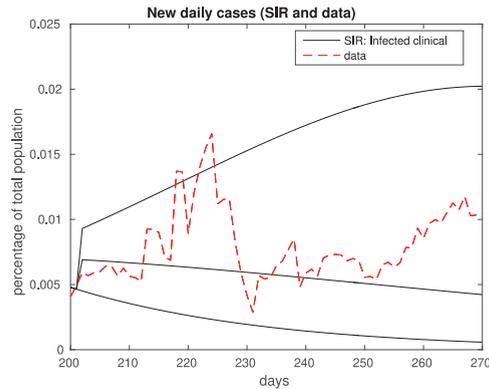[5] For instance, free market policies are discussed in https://johnhcochrane.blogspot.com/2020/12/free-market-vaccines.html.

**Fig. 7.** Number of daily new cases. The dashed red plot indicates the data and the solid black plot represents the simulation results for frozen $\beta_w$ and $\mu$ corresponding to the best/intermediary/worst case scenarios. Note that, for most of the 70-day forecast, that data lie within the interval predicted by the model.

Such heterogeneity leads to interesting immunization policy possibilities. People in the age group 15–19 (around 500 k people in Rio) have the most impact on virus dissemination, but few of them die. On the other hand, those above 60 (around 900 k people) are at high risk, but they have little impact on the spreading of the disease.

To offer a glimpse on these issues, we modify our effective contact matrix $C^e$, allowing for the possibility of immunization. Let us redefine (2) by

$$C^e(t) = V(t)\Big(\beta_h(t)C^{\text{home}} + \beta_w(t)C^{\text{work}} + \beta_s(t)C^{\text{school}} + \beta_o(t)C^{\text{other}}\Big)V(t), \tag{13}$$

where the transient diagonal matrix $V(\cdot)$ models the vaccinated age groups. Its $j$-th diagonal element is zero if the corresponding age group is protected, and one otherwise. Of course, it is also possible to consider partial immunization by setting values between zero and one.

Consider the hypothetical situation of having $\beta_h$ and $\beta_w$ as computed before (see Fig. 4) and the population of Rio. Between the days 200–500, we assume that $\beta_w = \beta_{opt} = 0.0064$ or $\beta_w = \beta_{max} = 0.0117$, and a constant lethality strength $\mu = 1$.

Suppose that there is a limited supply of vaccines, which are made available only at a certain moment, that is, the vaccination campaign takes place at a single day. In our examples, this is 100, 150 or 200 days after the pandemic starts. Of course, the timing of vaccination campaigns impacts its effectiveness, and there is an important body of literature investigating this issue (Albani et al., 2021b; Amaku et al., 2021a; Quach and Deeks, 2021).

To better compare experiments, we define an utility function that takes into account the number of lost lives. For $\Upsilon \in (0, 1)$, we define the $\Upsilon$-discounted cost of death as

$$\sum_{t=1}^{+\infty} \Upsilon^t d(t), \tag{14}$$

where $d(\cdot)$ is the number of daily fatalities. We consider $\Upsilon = 0.996^{1/365}$ in the numerical tests.

In our numerical experiments, we first assume that a single shot of the vaccine induces a perfect immunization and stop the virus spreading. In a second example, we consider partial immunization. Consider three age groups of similar size as targets of an immunization campaign: those aged 15–34 ($\approx 1,584$ people), 40–59 ($\approx 1,636$ people) and those above 50 ($\approx 1,670$ people). In all tested situations, the groups with best results were either the 15–34 or those above 50.

In Fig. 12, we plot the accumulated deaths considering full immunization of the groups aged 15–34 or over 50, for a fixed $\beta_w = 0.0064$ after day 200. In solid blue we plot the total deaths if no vaccination takes place. The remaining plots correspond to the total deaths if vaccination occur on day 100 (black), 150 (red) or 200 (magenta). The solid lines correspond to vaccinating the young and the dashes lines corresponding to vaccination of the elderly. In Fig. 13, we repeat the computations fixing a higher $\beta_w = 0.0117$ after day 200.

Basically, two different policies are possible: *stop the spreading* corresponds to vaccinating those in the 15–34 age group, and *protect the vulnerable* corresponds to vaccinating those above 50. Note from the results that the impact of stop the spreading decreases as the vaccination date is delayed. That is because the percentage of young population already contaminated at later dates is larger, and vaccinating the young group ends up not being so effective. Such effect is amplified for higher transmission rates, as one can see by comparing Figs. 12 and 13. Note also that for short terms, it is always preferable to protect the vulnerable, indicating that vaccinating the elderly population might be a wise option if a continuous supply of vaccines is available.
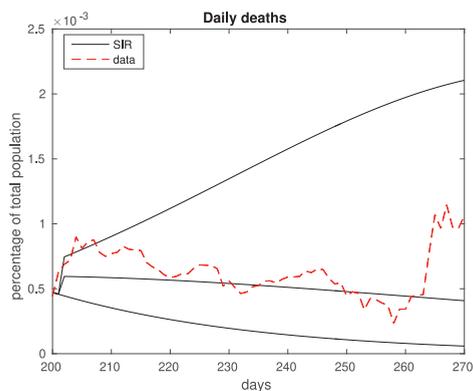
**Fig. 8.** Similarly to Fig. 7, the figure displays the number of daily deaths, dashed red being data and solid black being the best/intermediary/worst cases predicted by the model.
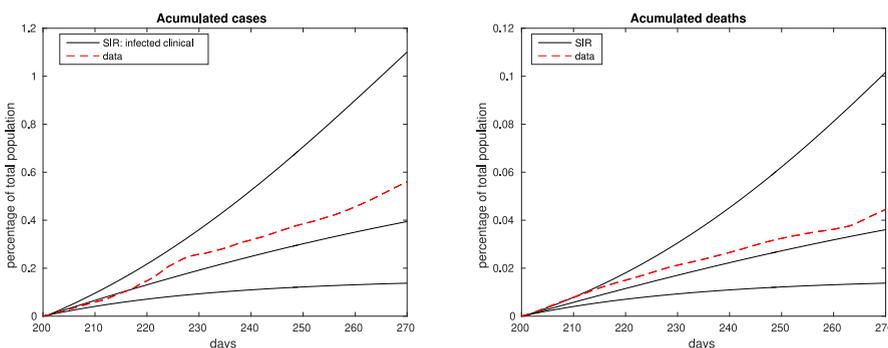


**Fig. 9.** We present here the accumulated number of cases (left) and deaths (right). In both figures, the dashed red plots correspond to real data and the solid black plots represent the simulation results for frozen $\beta_w$ and $\mu$, for the best/intermediary/worst case scenarios.
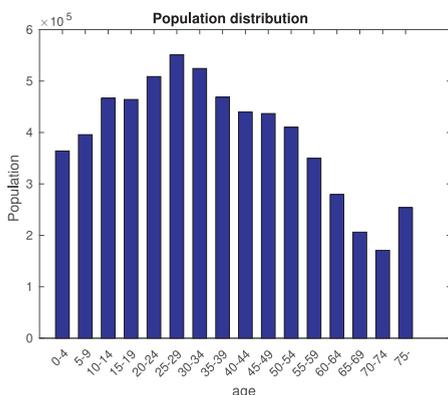


**Fig. 10.** Population profile of Rio. The last column corresponds to the total population above 75 years old.

Considering the total amount of casualties in the long run, for lower transmission rates it is always better to vaccinate the elders (Fig. 12). For higher transmission rates, these results are reversed (Fig. 13). However, when considering the cost of death (14), the results are mixed, as shown in Table 1. For related results, see (Meehan et al., 2020; 1093Giubilini, Savulescu, & Wilkinson, ; Zhao et al., 2020). We also remark that although most countries are pursuing policies that prioritize the elderly, there are exceptions (The Japan Times, 2020).

We consider in a second example the possibility that the vaccines are not perfect, and display the results in Fig. 14. Assume that a immunization campaign happens at day 150, and that either those aged 15–34 (solid lines) or those over 50 (dashed
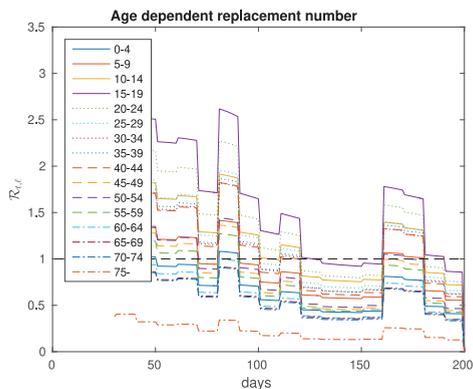
**Fig. 11.** Age-dependent replacement numbers corresponding to the dynamics of the disease in Rio.
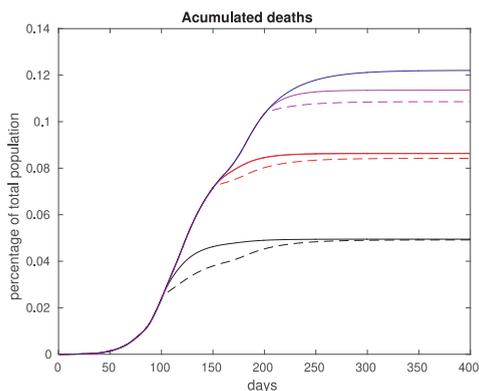


**Fig. 12.** Accumulated deaths for $\beta_w = 0.0064$. We assume no vaccination (solid blue), vaccination for group aged 15–34 (solid plots) and over 50 (dashed plots). The vaccination takes place on day 100 (black), 150 (red) or 200 (magenta). Note that vaccinating the elderly population is always optimal in the short term, but not necessarily in the long run.

plots) are vaccinated. Assume also that the immunization is perfect (in blue), or has partial efficiency of 70% (black) or 50% (red). Note that, as expected, as the vaccine efficiency decreases the number of casualties increase. It is maybe more surprising to realize that this effect is more aggravated for the elderly.
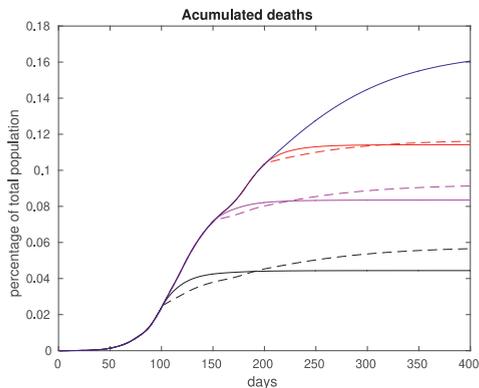


**Fig. 13.** Accumulated deaths for $\beta_w = 0.0117$, a higher rate of transmission. As before, we assume no vaccination (solid blue), vaccination for group aged 15–34 (solid plots) and over 50 (dashed plots). The vaccination takes place on day 100 (black), 200 (red) or 300 (magenta). Note that vaccinating the elderly population is always optimal in the short term, but not necessarily in the long run.

**Table 1**

Numerical investigation of immunization policies based on age. In each line we present the values of the death cost functional (14) assuming different age groups are vaccinated. The second column contains the population of the groups. In the three columns that follow, we present the results considering the vaccination takes place on the days 100, 150 or 200 and assuming that $\beta_w = 0.0064$. The following three columns display the results for $\beta_w = 0.0117$ and vaccination on the days 100, 150 or 200. For each column, we highlight the best results.

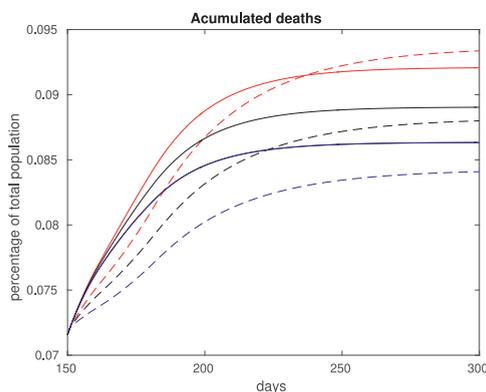| Age group | # vaccines | $\beta_w = 0.0064$ after day 200 | | | $\beta_w = 0.0117$ after day 200 | | |
| | | Vaccination day | | | | | |
| | | 100 | 150 | 200 | 100 | 150 | 200 |
|---|---|---|---|---|---|---|---|
| none | 0 | 7.4 k | | | 9.6 k | | |
| 15—34 | 1,584 k | **3 k** | 5.2 k | **5.9 k** | **3 k** | **5.3 k** | 7.1 k |
| 40—59 | 1,636 k | 3.3 k | 5.3 | 6.7 | 3.7 k | 5.7 k | 7.1 k |
| 50- | 1,670 k | **3 k** | **5.1 k** | 6.6 | 3.4 k | 5.5 k | **7 k** |



**Fig. 14.** Accumulated deaths for $\beta_w = 0.0064$, under varying vaccination efficiencies. Assume that at day 150, a vaccination campaign targets either the young (aged 15—34, solid plots) or those over 50 (dashed plots). The vaccination efficiency varies as follows: 100% (blue), 70% (black) or 50% (red).

## 5. Discussions and conclusion

Epidemics as the COVID-19 are hard to model, not only because of biological factors, but also due to the unpredictability of human responses. Political, economical, social and individual factors influence the behavior of people, leading to various degrees of risk behavior. Making long term predictions is riskier than in general (non-human) biological systems, but they are crucial for planning non-pharmaceutical interventions, allocation of resources, economical decisions, etc.

We propose in the work a general form to predict possible values of crucial parameters based on their history. Our method has two steps. Based on the dynamics of the disease on a multi-generational SIR model, we first determine the values of some time-dependent parameters, using a random optimization algorithm. Then, we feed the SIR model using extreme values of such parameters, and predict best/worst case scenarios. We are also consider a prediction that is in between the two extreme cases.

Our methodology is general and can be applied to other systems and circumstances, and might be useful to make long-term predictions in situations where unexpected occurrences might change the behavior of the system.

We also explore age-based immunization strategies, under simplifying hypotheses. For instance, while we assume the behavior of the population as known, the very existence of vaccines changes people's behavior. Moreover, several important aspects of vaccination are not considered here, as ethical problems, logistics, etc. We also point out that other policies might be better, as targeting certain classes of workers or individuals with comorbidities.

In our limited setting, we find out that optimal policies depend, among other things, on transmission rates, the vaccination efficiency and its timing. Earlier vaccination dates, low vaccine efficiency and higher transmission indices increase the impact of vaccinating the young. Under most conditions however, it pays off to vaccinate the elderly, although they contribute the least for spreading the viruses. Indeed, for short terms, protecting the older population is always preferred. That seems to indicate that under a continuous supply of vaccines, protecting the elderly is a wise option.

It is clear that the model provides different indications under different scenarios, and we shall explore more detailed policies in a future work.

As we point out in our Disclaimer, our goal is not to predict actual number of cases and deaths or to propose general policies, but instead to investigate limited aspects of the disease dynamics and raise possibilities of immunization strategies.

## Declaration of competing interest

## Acknowledgements

## References

d'Onofrio, A., & Manfredi, P. (2020). *The interplay between voluntary vaccination and reduction of risky behavior: A general behavior-implicit sir model for vaccine preventable infections*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-41120-6_10

Acemoglu, D., Chernozhukov, V., Werning, I., & Whinston, M. D. (May 2020). *A multi-risk sir model with optimally targeted lockdown*. National Bureau of Economic Research. https://doi.org/10.3386/w27102. Working Paper 27102 http://www.nber.org/papers/w27102.

Aguas, R., Corder, R. M., King, J. G., Goncalves, G., Ferreira, M. U., & Gomes, M. G. M.. Herd immunity thresholds for sars-cov-2 estimated from unfolding epidemics. medRxivarXiv https://www.medrxiv.org/content/early/2020/07/24/2020.07.23.20160762.full.pdf https://www.medrxiv.org/content/early/2020/07/24/2020.07.23.20160762.

Albani, V. V. L., Loria, J., Massad, E., & Zubelli, J. P. (2021b). *The impact of covid-19 vaccination delay: A modelling study for Chicago and NYC data*. arXiv:2102.12299.

Albani, V. V. L., Velho, R. M., & Zubelli, J. P. (2021a). Estimating, monitoring, and forecasting covid-19 epidemics: A spatiotemporal approach applied to nyc data. *Scientific Reports, 11*(1), 9089. https://doi.org/10.1038/s41598-021-88281-w

Alvarez, F. E., Argente, D., & Lippi, F. (April 2020). *A simple planning problem for COVID-19 lockdown*. National Bureau of Economic Research. https://doi.org/10.3386/w26981. Working Paper 26981 http://www.nber.org/papers/w26981.

Amaku, M., Covas, D. T., Bezerra Coutinho, F. A., Azevedo, R. S., & Massad, E.. Modelling the impact of delaying vaccination against sars-cov-2 assuming unlimited vaccines supply. medRxivarXiv:https://www.medrxiv.org/content/early/2021/02/23/2021.02.22.21252189.full.pdf https://www.medrxiv.org/content/early/2021/02/23/2021.02.22.21252189.

Amaku, M., Covas, D. T., Bezerra Coutinho, F. A., Azevedo Neto, R. S., Struchiner, C., Wilder-Smith, A., et al. (2021b). Modelling the test, trace and quarantine strategy to control the covid-19 epidemic in the state of São Paulo, Brazil. *Infectious Disease Modelling, 6*, 46–55. https://doi.org/10.1016/j.idm.2020.11.004. http://www.sciencedirect.com/science/article/pii/S2468042720300749

Amaral, F., Casaca, W., Oishi, C. M., & Cuminato, J. A. (2021). Towards providing effective data-driven responses to predict the Covid-19 in São Paulo and Brazil. *Sensors, 21*(2), 540. https://doi.org/10.3390/s21020540

Atkeson, A. (March 2020). *What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios, working paper 26867*. National Bureau of Economic Research. https://doi.org/10.3386/w26867. http://www.nber.org/papers/w26867

Bakhta, A., Boiveau, T., Maday, Y., & Mula, O. (2020). *Epidemiological short-term forecasting with model reduction of parametric compartmental models. application to the first pandemic wave of COVID-19 in France*. arXiv:2009.09200.

Barril, C., Calsina, Á., Cuadrado, S., & Ripoll, J.. On the basic reproduction number in continuously structured populations, Mathematical Methods in the Applied Sciences, 10.1002/mma.6787, n/a (n/a). arXiv https://onlinelibrary.wiley.com/doi/pdf/10.1002/mma.6787 https://onlinelibrary.wiley.com/doi/abs/10.1002/mma.6787.

Bendtsen Cano, O., Cano Morales, S., & Bendtsen, C.. COVID-19 modelling: The effects of social distancing. medRxivarXiv: https://www.medrxiv.org/content/early/2020/03/30/2020.03.29.20046870.full.pdf https://www.medrxiv.org/content/early/2020/03/30/2020.03.29.20046870.

Bertaglia, G., & Pareschi, L. (2020). *Hyperbolic models for the spread of epidemics on networks: Kinetic description and numerical methods*. arXiv:2007.04019.

Besse, C., & Faye, G. (2020). *Dynamics of epidemic spreading on connected graphs*. arXiv:2011.11996.

Bhopal, S. S., & Bhopal, R. (2020). Sex differential in covid-19 mortality varies markedly by age. *The Lancet, 396*(10250), 532–533. https://doi.org/10.1016/S0140-6736(20)31748-7. https://www.sciencedirect.com/science/article/pii/S0140673620317487

Borelli, L., & Góes, G. S. (2020). *A macroeconomia das epidemias: Heterogeneidade interestadual no brasil*. Preprint IPEA.

Brauer, F. (2017). Mathematical epidemiology: Past, present, and future. *Infectious Disease Modelling, 2*(2), 113–127. https://doi.org/10.1016/j.idm.2017.02.001. http://www.sciencedirect.com/science/article/pii/S2468042716300367

Brazeau, N. F., Verity, R., Jenks, S., Fu, H., Whittaker, C., Winskill, P., et al. (Oct. 2020). *COVID-19 infection fatality ratio: Estimates from seroprevalence*. Imperial College. report 34 https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-34-ifr/.

Britton, T., Ball, F., & Trapman, P.. A mathematical model reveals the influence of population heterogeneity on herd immunity to sars-cov-2. SciencearXiv https://science.sciencemag.org/content/early/2020/06/22/science.abc6810.full.pdf https://science.sciencemag.org/content/early/2020/06/22/science.abc6810.

Brotherhood, L., Kircher, P., Santos, C., & Tertilt, M. (May 2020). *An economic model of the COVID-19 epidemic: The importance of testing and age-specific policies*. IZA Institute of Labor Economics. report 13265 https://www.iza.org/publications/dp/13265.

Byambasuren, O., Cardona, M., Bell, K., Clark, J., McLaws, M.-L., & Glasziou, P. (2020). Estimating the extent of asymptomatic covid-19 and its potential for community transmission: Systematic review and meta-analysis. *Official Journal of the Association of Medical Microbiology and Infectious Disease Canada, 5*(4), 223–234. https://doi.org/10.3138/jammi-2020-0030. https://jammi.utpjournals.press/doi/10.3138/jammi-2020-0030

Calvetti, D., Hoover, A., Rose, J., & Somersalo, E. (2020). *Bayesian dynamical estimation of the parameters of an SE(A)IR COVID-19 spread model*. arXiv:2005.04365.

Chen, F., & Toxvaerd, F. (2014). The economics of vaccination. *Journal of Theoretical Biology, 363*, 105–117. https://doi.org/10.1016/j.jtbi.2014.08.003. http://www.sciencedirect.com/science/article/pii/S0022519314004512

Coelho, F. C., Carvalho, L. M., Lana, R. M., Cruz, O. G., Bastos, L. S., Codeço, C. T., et al.. Modeling the post-containment elimination of transmission of COVID-19. medRxivarXiv https://www.medrxiv.org/content/early/2020/06/17/2020.06.15.20132050.full.pdf https://www.medrxiv.org/content/early/2020/06/17/2020.06.15.20132050.

Cui, J., Zhang, Y., & Feng, Z. (2019). Influence of non-homogeneous mixing on final epidemic size in a meta-population model. *Journal of Biological Dynamics, 13*(sup1), 31–46. https://doi.org/10.1080/17513758.2018.1484186. pMID: 29909739. arXiv:10.1080/17513758.2018.1484186.

DeBiasi, R. L., & Delaney, M.. Symptomatic and asymptomatic viral shedding in pediatric patients infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Under the surface, JAMA Pediatrics. https://jamanetwork.com/journals/jamapediatrics/articlepdf/2770149/jamapediatrics_debiasi_2020_ed_200025_1598471468.36822.pdf. https://doi.org/10.1001/jamapediatrics.2020.3996.

Delamater, P. L., Street, E. J., Leslie, T. F., Yang, Y. T., & Jacobsen, K. H. (2019). Complexity of the basic reproduction number ($R_0$). *Emerging Infectious Diseases, 25*(1), 1–4. https://doi.org/10.3201/eid2501.171901. https://wwwnc.cdc.gov/eid/article/25/1/17-1901_article

Dennis, J., McGovern, A., Vollmer, S., & Mateen, B. A. (March toJune 2020). *Improving COVID-19 critical care mortality over time in England: A national cohort study*. https://doi.org/10.1101/2020.07.30.20165134. medRxivarXiv: https://www.medrxiv.org/content/early/2020/08/03/2020.07.30.20165134.full.pdf https://www.medrxiv.org/content/early/2020/08/03/2020.07.30.20165134.

Diekmann, O., Heesterbeek, J. A. P., & Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology, 28*(4), 365–382. https://doi.org/10.1007/BF00178324

Donnat, C., & Holmes, S. (2004). *Modeling the heterogeneity in COVID-19's reproductive number and its impact on predictive scenarios*, Article 05272. arXiv e-prints (2020) arXiv:2004.05272arXiv.

Eichenbaum, M. S., Rebelo, S., & Trabandt, M. (March 2020). *The macroeconomics of epidemics, Working Paper 26882*. National Bureau of Economic Research. https://doi.org/10.3386/w26882. http://www.nber.org/papers/w26882

Ferguson, N. M., Laydon, D., & Nedjati-Gilani, G.e. a. (03 2020). *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. Tech. rep. Imperial College London. https://doi.org/10.25561/77482

Funk, S., Bansal, S., Bauch, C. T., Eames, K. T., Edmunds, W. J., Galvani, A. P., et al. (2015). Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics, 10*, 21–25. https://doi.org/10.1016/j.epidem.2014.09.005. challenges in Modelling Infectious DIsease Dynamics http://www.sciencedirect.com/science/article/pii/S1755436514000541.

Giubilini, A., Savulescu, J., & Wilkinson, D.. COVID-19 vaccine: Vaccinate the young to protect the old?, Journal of Law and the Biosciences, 7 (1), lsaa050. arXiv: https://academic.oup.com/jlb/article-pdf/7/1/lsaa050/33744176/lsaa050.pdf https://doi.org/10.1093/jlb/lsaa050.

Gomes, M. G. M., Corder, R. M., King, J. G., Langwig, K. E., Souto-Maior, C., Carneiro, J., et al.. Individual variation in susceptibility or exposure to sars-cov-2 lowers the herd immunity threshold. medRxivarXiv: https://www.medrxiv.org/content/early/2020/05/21/2020.04.27.20081893.full.pdf https://www.medrxiv.org/content/early/2020/05/21/2020.04.27.20081893.

González, E., & Villena, M. J. (2020). On the spatial dynamics of vaccination: A spatial SIRS-V model. *Computers & Mathematics with Applications, 80*(5), 733–743. https://doi.org/10.1016/j.camwa.2020.04.015. http://www.sciencedirect.com/science/article/pii/S0898122120301620

Guerrieri, V., Lorenzoni, G., Straub, L., & Werning, I. (April 2020). *Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages?, Working Paper 26918*. National Bureau of Economic Research. https://doi.org/10.3386/w26918. http://www.nber.org/papers/w26918

Hamer, W. (1906). The milroy lectures on epidemic disease in england–the evidence of variability and of persistency of type. *The Lancet, 167*(4305), 569–574. https://doi.org/10.1016/S0140-6736(01)80187-2. originally published as Volume 1, Issue 4305 http://www.sciencedirect.com/science/article/pii/S0140673601801872.

Han, M. S., Choi, E. H., Chang, S. H., Jin, B.-L., Lee, E. J., Kim, B. N., et al.. Clinical characteristics and viral RNA detection in children with coronavirus disease 2019 in the Republic of Korea, JAMA Pediatrics, 10.1001/jamapediatrics.2020.3988 https://jamanetwork.com/journals/jamapediatrics/articlepdf/2770150/jamapediatrics_han_2020_oi_200069_1598471469.6651.pdf.

Havers, F. P., Reed, C., Lim, T., Montgomery, J. M., Klena, J. D., Hall, A. J., et al.. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020, JAMA internal Medicine. arXiv https://jamanetwork.com/journals/jamainternalmedicine/articlepdf/2768834/jamainternal_havers_2020_oi_200062.pdf https://doi.org/10.1001/jamainternmed.2020.4130.

Heneghan, C., Brassey, J., & Jefferson, T. (2020). COVID-19: What proportion are asymptomatic?. https://www.cebm.net/covid-19/covid-19-what-proportion-are-asymptomatic/. (Accessed 29 October 2020).

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review, 42*(4), 599–653. http://www.jstor.org/stable/2653135.

Horwitz, L., Jones, S. A., Cerfolio, R. J., Francois, F., Greco, J., Rudy, B., et al.. Trends in covid-19 risk-adjusted mortality rates in a single health system. medRxivarXiv: https://www.medrxiv.org/content/early/2020/08/14/2020.08.11.20172775.full.pdf https://www.medrxiv.org/content/early/2020/08/14/2020.08.11.20172775.

Iannelli, M., & Milner, F. (2017). The basic approach to age-structured population dynamics. In *Lecture notes on mathematical modelling in the life sciences*. Dordrecht: Springer https://doi.org/10.1007/978-94-024-1146-1. models, methods and numerics.

Jia, J., Ding, J., Liu, S., Liao, G., Li, J., Duan, B., et al. (2020). Modeling the control of COVID-19: Impact of policy interventions and meteorological factors. *Electronic Journal of Differential Equations, 2020*(23), 1–24.

Jin, J.-M., Bai, P., He, W., Wu, F., Liu, X.-F., Han, D.-M., et al. (2020). Gender differences in patients with COVID-19: Focus on severity and mortality. *Frontiers in Public Health, 8*, 152. https://doi.org/10.3389/fpubh.2020.00152. https://www.frontiersin.org/article/10.3389/fpubh.2020.00152

Jorion, P. (2006). *Value at risk: The new benchmark for managing financial risk*. McGraw-Hill.

Kermack, W. O., McKendrick, A. G., & Walker, G. T. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London - Series A: Containing Papers of a Mathematical and Physical Character, 115*(772), 700–721. https://doi.org/10.1098/rspa.1927.0118. https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1927.0118. https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0118

Kimathi, M., Mwalili, S., Ojiambo, V., & Gathungu, D. K. (2021). Age-structured model for COVID-19: Effectiveness of social distancing and contact reduction in Kenya. *Infectious Disease Modelling, 6*, 15–23. https://doi.org/10.1016/j.idm.2020.10.012. https://www.sciencedirect.com/science/article/pii/S246804272030066X

Klepac, P., & Caswell, H. (2011). The stage-structured epidemic: Linking disease and demography with a multi-state matrix approach model. *Theoretical Ecology, 4*(3), 301–319. https://doi.org/10.1007/s12080-010-0079-8

Klepac, P., Pomeroy, L. W., Bjørnstad, O. N., Kuiken, T., Osterhaus, A. D., & Rijks, J. M. (2009). Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak. *Proceedings of the Royal Society B: Biological Sciences, 276*(1666), 2469–2476. https://doi.org/10.1098/rspb.2009.0175. arXiv: https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.2009.0175 https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2009.0175.

Komatsu, B. K., & Menezes-Filho, N. (2020). *Simulando os efeitos de políticas de distanciamento social em São Paulo usando um modelo SEIR*. Maio.

Kremer, M. (1996). Integrating behavioral choice into epidemiological models of AIDS. *Quarterly Journal of Economics, 111*(2), 549–573. https://doi.org/10.2307/2946687. arXiv: https://academic.oup.com/qje/article-pdf/111/2/549/5460782/111-2-549.pdf.

Kuniya, T., & Wang, J. (2018). Global dynamics of an SIR epidemic model with nonlocal diffusion. *Nonlinear Analysis: Real World Applications, 43*, 262–282. https://doi.org/10.1016/j.nonrwa.2018.03.001. http://www.sciencedirect.com/science/article/pii/S1468121818301858

Lang, J. C., De Sterck, H., Kaiser, J. L., & Miller, J. C. (2018). Analytic models for SIR disease spread on random spatial networks. *Journal of Complex Networks, 6*(6), 948–970. https://doi.org/10.1093/comnet/cny004. arXiv: https://academic.oup.com/comnet/article-pdf/6/6/948/28007602/cny004.pdf.

Levin, A. T., Hanage, W. P., Owusu-Boaitey, N., Cochran, K. B., Walsh, S. P., & Meyerowitz-Katz, G. (2020). Assessing the age specificity of infection fatality rates for covid-19: Systematic review, meta-analysis, and public policy implications. *European Journal of Epidemiology, 35*(12), 1123–1138. https://doi.org/10.1007/s10654-020-00698-1

Li, X.-Z., Yang, J., & Martcheva, M. (2020). *Linear age-structured population models as a base of age-structured epidemic models*. Cham: Springer International Publishing.

Lourenço, J., Paton, R., Ghafari, M., Kraemer, M., Thompson, C., Simmonds, P., et al.. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the sars-cov-2 epidemic. medRxivarXiv: https://www.medrxiv.org/content/early/2020/03/26/2020.03.24.20042291.full.pdf https://www.medrxiv.org/content/early/2020/03/26/2020.03.24.20042291.

Manfredi, P., & d'Onofrio, A. (Eds.). (2013). *Modeling the interplay between human behavior and the spread of infectious diseases*. New York: Springer. https://doi.org/10.1007/978-1-4614-5474-8.

Matyaš, I. (1965). Random optimization. *Avtomat. i Telemeh, 26*, 246–253.

May, R. M., & Anderson, R. M. (1984). Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences, 72*(1), 83–111. https://doi.org/10.1016/0025-5564(84)90063-4

Meehan, M. T., Cocks, D. G., Caldwell, J. M., Trauer, J. M., Adekunle, A. I., Ragonnet, R. R., et al.. Age-targeted dose allocation can halve covid-19 vaccine requirements. medRxivarXiv: https://www.medrxiv.org/content/early/2020/12/02/2020.10.08.20208108.full.pdf https://www.medrxiv.org/content/early/2020/12/02/2020.10.08.20208108.

M'Kendrick, A. G. (1925). Applications of mathematics to medical problems. *Proceedings of the Edinburgh Mathematical Society, 44*, 98–130. https://doi.org/10.1017/S0013091500034428

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., et al. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine, 5*(3). https://doi.org/10.1371/journal.pmed.0050074, 1–1.

Paeng, S.-H., & Lee, J. (2017). Continuous and discrete sir-models with spatial distributions. *Journal of Mathematical Biology, 74*, 1709–1727. https://doi.org/10.1007/s00285-016-1071-8

Pastor-Satorras, R., & Vespignani, A. (2002). Immunization of complex networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics, 65*, Article 036104. https://doi.org/10.1103/PhysRevE.65.036104. https://link.aps.org/doi/10.1103/PhysRevE.65.036104

Peixoto, P. S., Marcondes, D., Peixoto, C., & Oliva, S. M. (2020). Modeling future spread of infections via mobile geolocation data and population dynamics. an application to COVID-19 in Brazil. *PloS One, 15*(7), 1–23. https://doi.org/10.1371/journal.pone.0235732

Perkins, T. A., & España, G. (2020). Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions. *Bulletin of Mathematical Biology, 82*(9), 118. https://doi.org/10.1007/s11538-020-00795-y

Perrings, C., Castillo-Chavez, C., Chowell, G., Daszak, P., Fenichel, E. P., Finnoff, D., et al. (2014). Merging economics and epidemiology to improve the prediction and management of infectious disease. *EcoHealth, 11*(4), 464–475. https://doi.org/10.1007/s10393-014-0963-6, 25233829[pmid] https://pubmed.ncbi.nlm.nih.gov/25233829.

Prem, K., Cook, A. R., & Jit, M. (2017). Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology, 13*(9), 1–21. https://doi.org/10.1371/journal.pcbi.1005697

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health, 5*(5), e261–e270. https://doi.org/10.1016/S2468-2667(20)30073-6

C. Quach, S. Deeks, Covid-19 vaccination: why extend the interval between doses?, Official Journal of the Association of Medical Microbiology and Infectious Disease Canadadoi:https://doi.org/10.3138/jammi-2021-0323. URL https://jammi.utpjournals.press/doi/full/10.3138/jammi-2021-0323.

Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling, 5*, 271–281. https://doi.org/10.1016/j.idm.2020.03.001. http://www.sciencedirect.com/science/article/pii/S2468042720300075

Sarma, M. S. (1990). On the convergence of the Baba and Dorea random optimization methods. *Journal of Optimization Theory and Applications, 66*(2), 337–343. https://doi.org/10.1007/BF00939542

Soofi, M., Najafi, F., & Karami-Matin, B. (2020). Using insights from behavioral economics to mitigate the spread of COVID-19. *Applied Health Economics and Health Policy, 18*(3), 345–350. https://doi.org/10.1007/s40258-020-00595-4

Sun, R. (2010). Global stability of the endemic equilibrium of multigroup sir models with nonlinear incidence. *Computers & Mathematics with Applications, 60*(8), 2286–2291. https://doi.org/10.1016/j.camwa.2010.08.020. http://www.sciencedirect.com/science/article/pii/S0898122110005845

Takács, B., & Hadjimichael, Y. (2019). *High order discretizations for spatial dependent SIR models.* arXiv:1909.01330.

The Japan Times. (2020). *With goal of herd immunity, Indonesia will give younger people vaccines first.* https://www.japantimes.co.jp/news/2020/12/16/asia-pacific/science-health-asia-pacific/indonesia-vaccine-young/.

Towers, S., & Feng, Z. (2012). Social contact patterns and control strategies for influenza in the elderly. *Mathematical Biosciences, 240*(2), 241–249. https://doi.org/10.1016/j.mbs.2012.07.007. http://www.sciencedirect.com/science/article/pii/S0025556412001721

Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., et al. (2020). Estimates of the severity of coronavirus disease 2019: A model-based analysis. *The Lancet Infectious Diseases, 20*(6), 669–677. https://doi.org/10.1016/S1473-3099(20)30243-7. https://www.sciencedirect.com/science/article/pii/S1473309920302437

Viguerie, A., Lorenzo, G., Auricchio, F., Baroli, D., Hughes, T. J. R., Patton, A., et al. (2020). *Simulating the spread of COVID-19 via spatially-resolved susceptible-exposed-infected-recovered-deceased (SEIRD) model with heterogeneous diffusion.* arXiv:2005.05320.

Walker, P. G. T., Whittaker, C., Watson, O. J., Baguelin, M., Winskill, P., Hamlet, A., et al. (2020). The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science, 369*(6502), 413–422. https://doi.org/10.1126/science.abc0035. arXiv:https://science.sciencemag.org/content/369/6502/413.full.pdf https://science.sciencemag.org/content/369/6502/413.

Walsh, K. A., Jordan, K., Clyne, B., Rohde, D., Drummond, L., Byrne, P., et al. (2020). Sars-cov-2 detection, viral load and infectivity over the course of an infection. *Journal of Infection, 81*(3), 357–371. https://doi.org/10.1016/j.jinf.2020.06.067. https://www.sciencedirect.com/science/article/pii/S0163445320304497

Wang, L., Wang, G., Li, X., Yu, S., Kim, M., Wang, Y., et al. (2021). Modeling and forecasting COVID-19. *Notices of the American Mathematical Society, 68*(4), 585–595. https://doi.org/10.1090/noti

Wen, X., Ji, J., & Li, B. (2018). Asymptotic profiles of the endemic equilibrium to a diffusive sis epidemic model with mass action infection mechanism. *Journal of Mathematical Analysis and Applications, 458*(1), 715–729. https://doi.org/10.1016/j.jmaa.2017.08.016. http://www.sciencedirect.com/science/article/pii/S0022247X17307734

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese center for disease control and prevention. *Journal of the American Medical Association, 323*(13), 1239–1242. https://doi.org/10.1001/jama.2020.2648. https://jamanetwork.com/journals/jama/articlepdf/2762130/jama_wu_2020_vp_200028.pdf

Wu, Y., & Zou, X. (2016). Asymptotic profiles of steady states for a diffusive SIS epidemic model with mass action infection mechanism. *Journal of Differential Equations, 261*(8), 4424–4447. https://doi.org/10.1016/j.jde.2016.06.028. http://www.sciencedirect.com/science/article/pii/S0022039616301644

Zhao, Z., Niu, Y., Luo, L., Hu, Q., Yang, T., Chu, M., et al. (2020). *The optimal vaccination strategy to control COVID-19: A modeling study based on the transmission scenario in Wuhan city, China.* SSRN Preprint https://ssrn.com/abstract=3719045.

Zhou, B., She, J., Wang, Y., & Ma, X. (2020). Duration of viral shedding of discharged patients with severe COVID-19. *Clinical Infectious Diseases, 71*(16), 2240–2242. https://doi.org/10.1093/cid/ciaa451. arXiv: https://academic.oup.com/cid/article-pdf/71/16/2240/34393006/ciaa451.pdf.