

Evaluating the Interplay between Transmissibility and Virulence of SARS-CoV-2 by Mathematical Modeling

L. P. LOMBARDI JUNIOR¹, H. M. YANG^{2*}, B. SPIRA³ and A. C. YANG⁴

Received on May 4, 2022 / Accepted on March 22, 2023

ABSTRACT. During the first months of 2020 SARS-CoV-2 spread to all continents, virtually reaching all countries. In the subsequent months, new variants emerged in different regions of the world. A mathematical model based on the Covid-19 natural history encompassing the age-dependent fatality was applied to evaluate SARS-CoV-2 transmissibility and virulence. Transmissibility was assessed by calculating the basic reproduction number R_0 and virulence by counting the proportion of severe Covid-19 cases and deaths. The model parameters were adjusted against the data observed in the state of São Paulo, Brazil, considering two different levels of virulence. The severe Covid-19 cases and deaths were three times higher and R_0 was 25% lower when the more virulent SARS-CoV-2 variant was compared to the less virulent variant. However, under the high-virulence scenario the number of transmitting individuals is 25% lower, mainly due to the isolation of symptomatic individuals. The corollary that transmission increases in the low virulence scenario is also true. The estimated parameters, using data from São Paulo up to May 13, 2020, showed that the Covid-19 epidemic predicted with low virulence SARS-CoV-2 transmission matched the observed data just before the beginning of the relaxation, which occurred by mid-June 2020. The assessment of the interplay between transmissibility can be applied to explain in somehow the appearance of gamma and omicron variants of concern in São Paulo.

Keywords: mathematical model, SARS-CoV-2, basic reproduction number, asymptomatic individuals, Covid-19 data.

1 INTRODUCTION

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), an RNA virus, spreads mostly through air by means of aerosols [1], but can be also transmitted by droplets produced by

*Corresponding author: Hyun Mo Yang hyunyang@ime.unicamp.br

¹Department of Applied Mathematics, State University of Campinas, Praça Sérgio Buarque de Holanda, 651, 13083-859 Campinas, SP, Brazil – E-mail: luispedro_jr@hotmail.com <https://orcid.org/0000-0001-8029-6662>

²Department of Applied Mathematics, State University of Campinas, Praça Sérgio Buarque de Holanda, 651, 13083-859 Campinas, SP, Brazil – E-mail: hyunyang@ime.unicamp.br <https://orcid.org/0000-0002-1711-363X>

³Department of Microbiology, Institute of Biomedical Science of University of São Paulo Av. Prof. Lineu Prestes, 1374, 05508-900 São Paulo, SP, Brazil – E-mail: benys@usp.br <https://orcid.org/0000-0001-5394-9074>

⁴Division of Allergy and Immunology, General Hospital of the Medicine School of University of São Paulo, Av. Dr. Eneas Carvalho de Aguiar, 255, 05403-000 São Paulo, SP, Brazil – E-mail: arianacy@gmail.com <https://orcid.org/0000-0002-1901-3472>

coughing, sneezing or talking, or by fomites [2] [3]. The coronavirus disease 2019 (Covid-19) caused by SARS-CoV-2 infection was declared a pandemic by the World Health Organization on March 11, 2020. In general, the fatality rate is considerably higher in senior patients (60 years or more) [4].

SARS-CoV-2 is an RNA-virus presenting relatively high mutation rate (See Appendix A for details). Briefly, the original SARS-CoV-2 variant was first detected in Wuhan, China by December 2019, and the omicron variant, that appeared in South Africa in November 2021, is the most phylogenetically divergent variant of concern. The omicron variant showed high transmissibility but less virulence than the original SARS-CoV-2.

We developed a mathematical model to evaluate the interplay between transmissibility and virulence during the Covid-19 epidemic in the state of São Paulo, Brazil. We hypothesized that transmissibility is measured by the transmission rate (ultimately, by the basic reproduction number R_0), and that virulence is assessed by the ratio between asymptomatic and symptomatic individuals. It is worth stressing that the additional mortality due to Covid-19 results in an asymptotic value, which is not an equilibrium value at steady state. Hence, a special method is required to assess the steady state to obtain R_0 . Additionally, as individuals with severe Covid-19 are hospitalized (isolated), and they must not be considered as transmitting agents of SARS-CoV-2, rather only the asymptomatic or pre-symptomatic or mild Covid-19 individuals are spreading the disease [19]. In section 2, we present and analyze a model that describes SARS-CoV-2 transmission, in section 3, using the data for São Paulo State, we estimate the basic reproduction number R_0 considering two different virulence levels. The discussion and conclusion are given in sections 4 and 5.

2 MATERIAL AND METHODS

One of the main aspects of Covid-19 is the high fatality rate of older people. For this reason, we divided the population in two groups, one composed of young (60 years old or less, denoted by subscript y) and the other containing individuals of 60 years old or more, denoted by subscript o . The community's vital dynamic is described by the per-capita rates of birth (ϕ) and mortality (μ), and φ is the aging rate, that is, the flow from the young subpopulation y to the senior subpopulation o . Another aspect is SARS-CoV-2 transmission by asymptomatic and pre-symptomatic individuals, that is, individuals that never develop symptoms and those before the onset of the disease [20]. Hence, for each subpopulation j ($j = y, o$), individuals are divided into seven compartments: susceptibles S_j , isolated (quarantine) Q_j , exposed and incubating E_j , asymptomatic A_j , pre-symptomatic (or pre-diseased) individuals P_j , symptomatic individuals with mild Covid-19 M_j , and individuals with severe Covid-19 D_j . All young and senior individuals in classes A_j , M_j , and D_j enter to the same recovered class R . Hence, the SQAEMDR model has 15 compartments. The compartments irrespective to age are $E = E_y + E_o$ (exposed), $A = A_y + A_o$ (asymptomatic), $P = P_y + P_o$ (pre-symptomatic), $M = M_y + M_o$ (mild Covid-19), and $D = D_y + D_o$ (severe Covid-19). Table 1 summarizes the model compartments (variables).

Table 1: Summary of the model variables, for $j = y$ (young) and $j = o$ (senior).

Symbol	Meaning
S_j	Susceptible individuals
Q_j	Isolated individuals
E_j	Exposed and incubating SARS-CoV-2 individuals
A_j	Asymptomatic individuals
P_j	Pre-diseased (pre-symptomatic) individuals
M_j	Mild (non-hospitalized) Covid-19 individuals
D_j	Severe (hospitalized) Covid-19 individuals
R	Recovered (immune) individuals

The susceptible individuals are either in isolation (quarantine) or circulating. To express isolation (the flow from S to Q), we consider a unique pulse in isolation at time $t = \tau$, described by $u_j S_j \delta(t - \tau)$, with $j = y, o$ – The fraction of isolated people is u_j and $\delta(x)$ is the Dirac delta function, that is, $\delta(x) = \infty$, if $x = 0$, otherwise, $\delta(x) = 0$, with $\int_0^\infty \delta(x) dx = 1$. The equations that describe isolation are as follows

$$\begin{cases} \frac{d}{dt} S_y &= \phi N - (\phi + \mu) S_y - \lambda S_y - u_y S_y \delta(t - \tau) \\ \frac{d}{dt} S_o &= \phi S_y - \mu S_o - \lambda \psi S_o - u_o S_o \delta(t - \tau) \\ \frac{d}{dt} Q_j &= u_j S_j \delta(t - \tau) - \mu Q_j. \end{cases} \tag{2.1}$$

On the other hand, the natural history of Covid-19 is the same for the young ($j = y$) and senior ($j = o$) subpopulations in circulation. We assume that individuals in the asymptomatic (A_j), pre-symptomatic (P_j), and a fraction z_j of mild Covid-19 (M_j) class are transmitting the virus. Other infected classes ($(1 - z_j) M_j$ and D_j) are under voluntary or forced isolation. Susceptible individuals are infected at a rate $\lambda_j S_j$ (known as the mass action law [21]), where λ_j is the per-capita incidence rate (or force of infection) defined by $\lambda_j = \lambda (\delta_{jy} + \psi \delta_{jo})$, with λ being

$$\lambda = \frac{\epsilon}{N} (\beta_{1y} A_y + \beta_{2y} P_y + \beta_{3y} z_y M_y + \beta_{1o} A_o + \beta_{2o} P_o + \beta_{3o} z_o M_o), \tag{2.2}$$

where δ_{ij} is the Kronecker delta, with $\delta_{ij} = 1$ if $i = j$, and 0, if $i \neq j$; and β_{1j} , β_{2j} and β_{3j} are the transmission rates – the rates at which a virus encounters a susceptible people and infects him/her. The parameter $\psi \geq 1$ measures more susceptibility among seniors than youngs. The parameter $\epsilon \leq 1$ diminishes the transmission rates – the protection factor ϵ decreases viral transmission by individual (face mask, hygiene, etc.) and collective (social distancing) protective measures.

Susceptible individuals are infected at rate λ_j and enter into class E_j . After an average period $1/\sigma_j$ in class E_j , where σ_j is the incubation rate, exposed individuals develop into asymptomatic A_j (with probability l_j) or pre-symptomatic class P_j (with probability $1 - l_j$). After an average period $1/\gamma_j$ in class A_j , where γ_j is the recovery rate of asymptomatic individuals, asymptomatic

individuals acquire immunity (recovered) and enter the recovered class R . Asymptomatic individuals may manifest symptoms at the end of this period, and a fraction $1 - \chi_j$ becomes mildly ill (class M_j). As for the symptomatic individuals, after an average period $1/\gamma_{1j}$ in class P_j , where γ_{1j} is the infection rate of pre-symptomatic individuals, these individuals become either severely ill (class D_j) (with probability $1 - k_j$) or mildly ill (class M_j) (with probability k_j). Individuals in class D_j acquire immunity after a period $1/\gamma_{2j}$, where γ_{2j} is the recovery rate of severe Covid-19, and enter the recovered class R or die under the disease-induced (additional) mortality rate α_j . Individuals in the mild Covid-19 class M_j acquire immunity after a period $1/\gamma_{3j}$, where γ_{3j} is the recovery rate of mild Covid-19, and enter the recovered class R . Table 2 summarizes the model parameters. The description of the assigned values to the Covid-19 natural history parameters can be found in [22] (the parameters marked with * are estimated considering two SARS-CoV-2 virulence levels).

Table 2: Summary of the model parameters ($j = y, o$) and values (rates in $days^{-1}$ and proportions are dimensionless) from [22]. (*) Values fitted considering two levels of the SARS-CoV-2 virulence.

Symbol	Meaning	Value
μ	Natural mortality rate	$1/(78.4 \times 365)$
ϕ	Birth rate	$1/(78.4 \times 365)$
φ	Aging rate	6.7×10^{-6}
$\sigma_y (\sigma_o)$	Incubation rate	$1/5.8 (1/5.8)$
$\gamma_y (\gamma_o)$	Recovery rate of asymptomatic individuals	$1/12 (1/14)$
$\gamma_{1y} (\gamma_{1o})$	Infection rate of pre-symptomatic individuals	$1/4 (1/4)$
$\gamma_{2y} (\gamma_{2o})$	Recovery rate of severe Covid-19 individuals	$1/12 (1/21)$
$\gamma_{3y} (\gamma_{3o})$	Infection rate of mild Covid-19 individuals	$1/13 (1/16)$
τ	Time of the introduction of isolation	March 24, 2020
$z_y (z_o)$	Proportion of transmission by mild Covid-19 individuals	$0.5 (0.2)$
ψ	Increased susceptibility among seniors	1.17
$\chi_y (\chi_o)$	Proportion of remaining as asymptomatic individuals	$0.98 (0.95)$
$l_y (l_o)$	Proportion of asymptomatic individuals	*
$k_y (k_o)$	Proportion of mild (non-hospitalized) Covid-19	*
ε	Protection factor	*
$u_y (u_o)$	Proportion in isolation	*
$\alpha_y (\alpha_o)$	Additional mortality rate	*
$\beta_{1y} (\beta_{1o})$	Transmission rate due to asymptomatic individuals	*
$\beta_{2y} (\beta_{2o})$	Transmission rate due to pre-symptomatic individuals	*
$\beta_{3y} (\beta_{3o})$	Transmission rate due to mild Covid-19 individuals	*

The SARS-CoV-2 transmission model is described by the system of ordinary differential equations, named the SQEAPMDR model considering young (y) and senior (o) subpopulations. The dynamic equations for S , Q , E , A , P , M , D , and R are obtained through the balance of inflow

and outflow in each compartment. We drop out the pulses in equation (2.1) describing quarantine individuals and transfer them to the pulse conditions at time of quarantine. Hence, the equations for the susceptible individuals are

$$\begin{cases} \frac{d}{dt}S_y &= \phi N - (\phi + \mu)S_y - \lambda S_y \\ \frac{d}{dt}S_o &= \phi S_y - \mu S_o - \lambda \psi S_o, \end{cases} \tag{2.3}$$

for isolated individuals Q_j and infected individuals, with $j = y, o$,

$$\begin{cases} \frac{d}{dt}Q_j &= -\mu Q_j \\ \frac{d}{dt}E_j &= \lambda (\delta_{jy} + \psi \delta_{jo})S_j - (\sigma_j + \mu)E_j \\ \frac{d}{dt}A_j &= l_j \sigma_j E_j - (\gamma_j + \mu)A_j \\ \frac{d}{dt}P_j &= (1 - l_j) \sigma_j E_j - (\gamma_{1j} + \mu)P_j \\ \frac{d}{dt}M_j &= (1 - \chi_j) \gamma_j A_j + k_j \gamma_{1j} P_j - (\gamma_{3j} + \mu)M_j \\ \frac{d}{dt}D_j &= (1 - k_j) \gamma_{1j} P_j - (\gamma_{2j} + \mu + \alpha_j)D_j, \end{cases} \tag{2.4}$$

and for recovered individuals,

$$\frac{d}{dt}R = \chi_y \gamma_y A_y + \gamma_{3y} M_y + \gamma_{2y} D_y + \chi_o \gamma_o A_o + \gamma_{3o} M_o + \gamma_{2o} D_o - \mu R, \tag{2.5}$$

where $N_j = S_j + Q_j + E_j + A_j + P_j + M_j + D_j$, and $N = N_y + N_o + R$ obeys

$$\frac{d}{dt}N = (\phi - \mu)N - \alpha_y D_y - \alpha_o D_o, \tag{2.6}$$

with the initial number of population being $N(0) = N_0 = N_{0y} + N_{0o}$, where N_{0y} and N_{0o} are the size of the young and senior subpopulations at $t = 0$.

The initial conditions (simulation time $t = 0$ corresponding to the calendar time when the first case was confirmed: February 26 for São Paulo State) supplied to the system of equations (2.3), (2.4), and (2.5) are, for $j = y, o$,

$$S_j(0) = N_{0j}, \quad X_j(0) = n_{X_j}, \quad \text{where } X_j = Q_j, E_j, A_j, P_j, M_j, D_j, R, \tag{2.7}$$

where n_{X_j} is a non-negative number. For instance, $n_{E_y} = n_{E_o} = 0$ implies that there was not any exposed person (young or senior) at the beginning of the epidemic. The isolation implemented at $\tau = 27$ (corresponding to calendar time March 24, 2020) is described by the pulse conditions at time τ

$$\begin{cases} S_j(\tau^+) = S_j(\tau^-)(1 - u_j) & \text{and} & Q_j(\tau^+) = Q_j(\tau^-) + S_j(\tau^-)u_j, \\ X_j(\tau^+) = X_j(\tau^-), & \text{where} & X_j = E_j, A_j, P_j, M_j, D_j, R, \end{cases} \tag{2.8}$$

with $\tau^- = \lim_{t \rightarrow \tau^-} t$ (for $t < \tau$), and $\tau^+ = \lim_{t \rightarrow \tau^+} t$ (for $t > \tau$). For São Paulo State, at $t = 0$ we use $N_{0y} = 37.8$ million and $N_{0o} = 6.8$ million [36].

In this paper, we use, at $t = 0$,

$$\left\{ \begin{array}{l} \left(S_y(0) = N_{0y}, \quad E_y(0) = 30, \quad A_y(0) = 24, \quad P_y(0) = 6, \quad M_y(0) = 6, \quad D_y(0) = 0 \right) \\ \text{and} \\ \left(S_o(0) = N_{0o}, \quad E_o(0) = 20, \quad A_o(0) = 16, \quad P_o(0) = 4, \quad M_o(0) = 3, \quad D_o(0) = 1, \right) \end{array} \right. \quad (2.9)$$

plus $Q_y(0) = Q_o(0) = R(0) = 0$. The initial conditions' setup are found in [22]. Briefly, at $t = 0$, one senior individual was diagnosed with severe Covid-19. For other variables, using the ratio 4 : 1 for asymptomatic:pre-symptomatic and 3 : 1 for mild:severe (non-hospitalized:hospitalized) Covid-19, if we assume that there is one person in D_o (the first confirmed case), then there are 3 persons in M_o . The sum 4 is the number of persons in class P_o , implying that there are 16 in class A_o , and the sum 20 is the number of persons in class E_o . For young population, we used $D_y(0) = 0$ and arbitrarily $M_y(0) = 6$. Finally, at the beginning of the epidemic, there are not isolated and immunized persons.

The system of equations (2.3), (2.4), and (2.5) result in an asymptote, which is not an equilibrium value (or steady state of the system) due to the additional equation (2.6). In Appendix B, the steady-state of the system of equations in terms of fractions corresponding to equations (2.3), (2.4) and (2.5) was analyzed to obtain the basic reproduction number R_0 , which is given by equation (B.5) in Appendix B. We consider this equation as an approximate value to the time-varying population model, and write the fractions as $s_y^0 = N_{0y}/N$ and $s_o^0 = N_{0o}/N$, resulting in

$$R_0 = R_{0y} + R_{0o}, \quad \text{with} \quad R_{0y} = (R_{1y} + R_{2y}) \frac{N_{0y}}{N_0} \quad \text{and} \quad R_{0o} = (R_{1o} + R_{2o}) \frac{N_{0o}}{N_0}, \quad (2.10)$$

where R_{0y} and R_{0o} are the partial reproduction numbers for young and senior subpopulations, and N_{0y} and N_{0o} are the initial numbers for young and senior subpopulations, with $N_0 = N_{0y} + N_{0o}$, and

$$\left\{ \begin{array}{l} R_{1y} = l_y \frac{\sigma_y}{\sigma_y + \phi} \frac{\beta_{1y}}{\gamma_y + \phi} + (1 - l_y) \frac{\sigma_y}{\sigma_y + \phi} \frac{\beta_{2y}}{\gamma_{1y} + \phi} \\ R_{1o} = l_o \frac{\sigma_o}{\sigma_o + \phi} \frac{\beta_{1o}\Psi}{\gamma_o + \phi} + (1 - l_o) \frac{\sigma_o}{\sigma_o + \phi} \frac{\beta_{2o}\Psi}{\gamma_{1o} + \phi} \\ R_{2y} = \left[l_y \frac{\sigma_y}{\sigma_y + \phi} (1 - \chi_y) \frac{\gamma_y}{\gamma_y + \phi} + (1 - l_y) \frac{\sigma_y}{\sigma_y + \phi} k_y \frac{\gamma_{1y}}{\gamma_{1y} + \phi} \right] \frac{z_y \beta_{3y}}{\gamma_{3y} + \phi} \\ R_{2o} = \left[l_o \frac{\sigma_o}{\sigma_o + \phi} (1 - \chi_o) \frac{\gamma_o}{\gamma_o + \phi} + (1 - l_o) \frac{\sigma_o}{\sigma_o + \phi} k_o \frac{\gamma_{1o}}{\gamma_{1o} + \phi} \right] \frac{z_o \beta_{3o}\Psi}{\gamma_{3o} + \phi} \end{array} \right. \quad (2.11)$$

Letting $z_y = z_o = 0$ ($R_{2y} = R_{2o} = 0$), we retrieve the basic reproduction number obtained in [23].

We chose two key parameters to evaluate the interplay between transmissibility and virulence.

- (1) **Transmissibility** – The force of infection λ given by equation (2.2) is the per-capita incidence rate, and λS_y (λS_o) is the total number of the new cases in the young (or senior) subpopulation per unit of time. Notwithstanding, the intensity of infection is proportional to the amount of virus released by infectious individuals and the capacity to infect susceptible individuals. The transmission rates depend on other characteristics not considered

here (social network, movement, demography, genetic, nutritional and health conditions, etc.). For this reason, we assume that the basic reproduction number R_0 calculated by the estimated transmission rates is an approximate measure for the transmissibility of SARS-CoV-2. To estimate the transmission rates β_{1j} , β_{2j} , and β_{3j} , $j = y, o$, the proportion of individuals in isolation u , and the protection factor ε , we used the accumulated number of severe Covid-19 cases $\Omega = \Omega_y + \Omega_o$ given by

$$\frac{d}{dt}\Omega_y = (1 - k_y)\gamma_{1y}P_y \quad \text{and} \quad \frac{d}{dt}\Omega_o = (1 - k_o)\gamma_{1o}P_o, \quad \text{with} \quad \Omega_y(0) = 0 \quad \text{and} \quad \Omega_o(0) = 0, \quad (2.12)$$

which represents the moving from class P to class D .

- (2) **Virulence** – By enhancing the virus' capacity to infect target cells, an increased number of cells are infected. Consequently, the possibility of manifesting mild or severe disease increases, increasing the number of symptomatic Covid-19 cases. For this reason, we equate virulence with the ratio between asymptomatic and symptomatic individuals. It is worth noticing that another effect of increased virulence is the high amount of virus released by infectious individuals in the environment, which is likely to increase transmission [8]. We consider two broadly separate values to express the proportion of asymptomatic individuals and estimate the transmission rates to calculate the basic reproduction number R_0 .

Besides the severe Covid-19 cases, the number of Covid-19 fatalities are registered. Based on these data, we can estimate the additional mortality rates α_j , $j = y, o$, using the number of deaths due to severe Covid-19 cases given by $\Pi = \Pi_y + \Pi_o$, where

$$\frac{d}{dt}\Pi_y = \alpha_y D_y \quad \text{and} \quad \frac{d}{dt}\Pi_o = \alpha_o D_o, \quad \text{with} \quad \Pi_y(0) = 0 \quad \text{and} \quad \Pi_o(0) = 0. \quad (2.13)$$

In the estimation of additional mortality rates, we must bear in mind that the time at which new cases and deaths were registered does not have direct correspondence; instead, they are delayed by Δ days, that is, $\Pi(t + \Delta) = \alpha D(t)$. From those individuals leaving compartment D (γ_j , μ , and α_j), we are counted only the deaths caused by severe Covid-19.

3 RESULTS

We studied the transmissibility and virulence of SARS-CoV-2 in the state of São Paulo, Brazil [24]. At the beginning of the epidemic, we had two data sets: severe Covid-19 cases (those in hospitals who were tested and confirmed) and deaths. During the period of the lack of mass testing (PCR and serology), the epidemic curve of SARS-CoV-2 can be characterized by the severe Covid-19 curve D .

To evaluate the interplay between transmissibility and virulence, we assumed that the basic reproduction number R_0 measures transmissibility, and the proportions of asymptomatic individuals l_y and l_o represent virulence. Letting $l = l_y = l_o$, we consider two values: $l_1 = 0.8$ (lower virulence, with the ratio between asymptomatic and symptomatic individuals being 4 : 1) and $l_2 = 0.2$ (higher virulence, with the ratio 1 : 4).

To estimate the transmission rates, the proportion of individuals in isolation, and the protective factor (x stands for one of these parameters), we calculate

$$\text{Sum} = \sum_{i=1}^n \left[\Omega(x, t_i) - \Omega^{ob}(t_i) \right]^2,$$

where $\Omega(x, t_i)$ is the accumulated severe Covid-19 cases calculated from equation (2.12) and $\Omega^{ob}(t_i)$ is the accumulated severe Covid-19 registered cases at day t_i . We searched for the value of x by minimizing the Sum. To evaluate Ω , the numerical solutions of equations (2.3), (2.4) and (2.5) are obtained using the initial conditions given by equation (2.9). To estimate the additional mortality rates, we substitute $\Omega(x, t_i)$ by $\Pi(\alpha, t_i)$ given by equation (2.13) and $\Omega^{ob}(t_i)$ by $\Pi^{ob}(t_i)$.

The observed data from February 26 to November 30 [24] is partitioned into two sets – The first set is used to estimate the model parameters (input data set), and the estimated model is then confronted with the second set to assess its prediction ability (test data set) [25]. We estimate the model parameters using severe Covid-19 (Ω^{ob}) and deaths (Π^{ob}) data from February 26 to May 13, 2020 (input data set). Then, we use the estimated parameters to predict the epidemic under interventions (partial quarantine and protective measures) from May 14 to November 30 (test data set) and compared the outcomes with the observed Ω^{ob} and Π^{ob} data.

3.1 Input data set – February 26 to May 13, 2020

We estimated the model parameters using the severe Covid-19 data Ω^{ob} from February 26 to May 13 following the estimation procedure presented in [22]. Briefly, the input data set was split into three periods aiming the estimation of the basic reproduction number (without any interventions), the proportion isolated by quarantine, and the effect of protective measures. For low virulence ($l_1 = 0.8$), we borrowed the estimated parameters from [22]. For high virulence ($l_2 = 0.2$), we obtain:

- (1) **Transmission rates** (data from February 26 to April 3) – Letting $u = 0$ and $\varepsilon = 1$, the estimated values are $\beta_y = 0.69$ and $\beta_o = 0.79$ (both in $days^{-1}$), resulting in the basic reproduction number $R_0 = 7.02$ using equation (2.10). The partial reproduction numbers are $R_{0y} = 6.14$ and $R_{0o} = 0.88$. This period, at the beginning of the Covid-19 epidemic without any kind interventions, characterizes the natural epidemic, allowing the estimation of the basic reproduction number.
- (2) **Proportion of individuals in partial quarantine** (data from March 24 to April 13) – Using the previously estimated β_y and β_o , and letting $\varepsilon = 1$, the estimated value is $u = 0.48$.
- (3) **Protective measures** (data from April 4 to May 13) – Using the previously estimated β_y , β_o , and u , the estimated value is $\varepsilon = 0.53$.

Considering the previously estimated parameters β_y , β_o , u , and ε based on Ω^{ob} , the Covid-19 fatality data Π^{ob} allow us to estimate:

- (4) **Additional mortality rates** (data from March 16 to May 28) – The estimated values for high virulence are $\alpha_y = 0.0018$ and $\alpha_o = 0.0071$ (both in $days^{-1}$) letting $\Delta = 15$ days. The first death due to Covid-19 was on March 16, and we considered 15 more observed data in the input data set due to Δ . The estimates for low virulence are found in [22].

The estimated parameters for $l_1 = 0.8$ and $l_2 = 0.2$ are summarized in Table 3. The probability of seniors becoming mildly ill k_j is lower than youngs [37]. The ratio R_{0o}/R_{0y} is 0.195 for low, and 0.143 for high virulent variants.

Table 3: The estimated parameters ($j = y, o$) considering two SARS-CoV-2 virulence levels (rates in $days^{-1}$ and proportions are dimensionless).

Parameter	Low ($l_1 = 0.8$) [22]	High ($l_2 = 0.2$)
$k_y (k_o)$	0.92 (0.75)	0.94 (0.78)
ϵ	0.5	0.53
$u_y (u_o)$	0.53 (0.53)	0.48 (0.48)
$\alpha_y (\alpha_o)$	0.00185 (0.0071)	0.00185 (0.0071)
$\beta_{1y} (\beta_{1o})$	0.78 (0.90)	0.69 (0.79)
$\beta_{2y} (\beta_{2o})$	0.78 (0.90)	0.69 (0.79)
$\beta_{3y} (\beta_{3o})$	0.78 (0.90)	0.69 (0.79)
$R_{0y} (R_{0o}) [R_0]$	7.73 (1.51) [9.24]	6.14 (0.88) [7.02]

Estimations of the epidemic parameters were done considering São Paulo state’s data from February 26 to May 13, 2020, for Ω ; and from March 16 to May 28, 2020, for Π . Notice that since April 4, the observed data correspond to the epidemic with interventions (quarantine u and protective measures ϵ).

Figure 1 shows the estimated curve of accumulated Covid-19 cases Ω and the observed data Ω^{ob} (a), and the accumulated deaths Π and the observed Π^{ob} (b) for the high (dashed line) and low (continuous line) virulence. Both low and high virulence levels provided quite similar estimated curves for the input data set.

To assess the contributions of young and senior subpopulations in the overall epidemic, Figure 2 shows the estimated curve of accumulated Covid-19 cases $\Omega_y, \Omega_o,$ and Ω (a), and the accumulated deaths Π_y, Π_o, Π (b) for the high (dashed line) and low (continuous line) virulence. Notice that $\Omega = \Omega_y + \Omega_o$ and $\Pi = \Pi_y + \Pi_o$ are those shown in Figure 1. For low virulence (high virulence), the numbers of severe Covid-19 cases on May 13 are 35408 (32251) and 22719 (24475) for young and senior subpopulations, with the ratio being 1.56 (1.32). For low virulence (high virulence), the numbers of Covid-19 fatality cases on May 28 are 988 (910) and 3123 (3406) for young and senior subpopulations, with the ratio being 0.32 (0.27).

As shown in Figures 1 and 2, both low and high virulence levels of SARS-CoV-2 provided a very good estimation of the epidemic under interventions. We must, however, choose one of them to

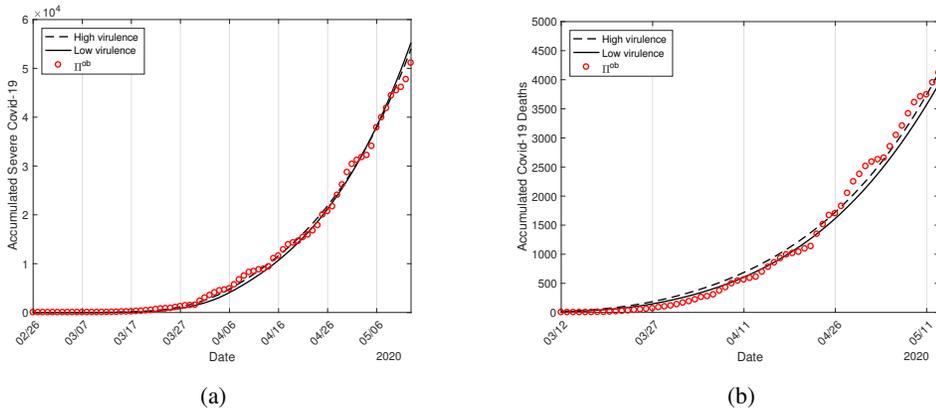


Figure 1: The estimated accumulated curve Ω and the observed data Ω^{ob} (a), and the accumulated deaths Π and the observed data Π^{ob} (b) for the higher (dashed line) and lower (continuous line) virulence. The natural epidemic corresponding to the low virulence is shown. Estimations were done considering data from São Paulo State.

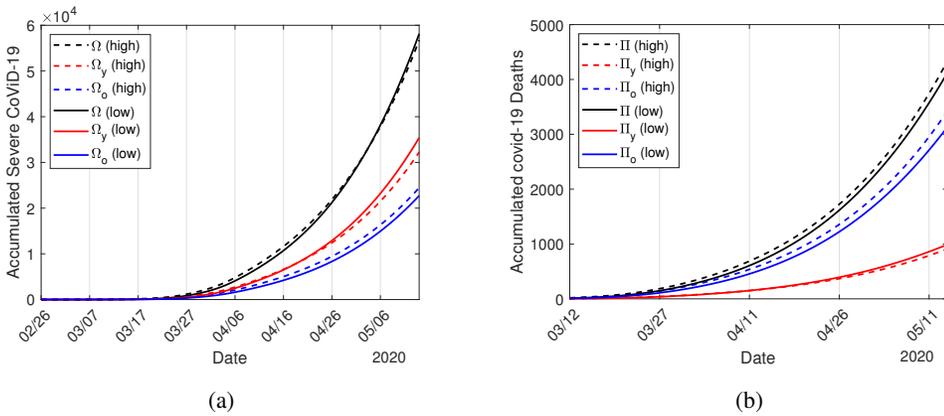


Figure 2: The estimated accumulated curves Ω_y , Ω_o , and Ω (a), and the accumulated deaths Π_y , Π_o , and Π (b) for the higher (dashed line) and lower (continuous line) virulence.

characterize the Covid-19 epidemic in São Paulo, which task is addressed by using the test data set.

3.2 Test data set – May 14 to November 30, 2020

We aim the evaluation of the estimated parameters’ prediction ability for two different virulence levels. The model is solved numerically during the test data sets using the parameters given in Table 2 and those fitted using the input data set given in Table 3. The results are compared with the observed data Ω^{ob} and Π^{ob} from São Paulo.

Figure 3 shows the extension of Figure 1 until November 30, 2020: the accumulated Covid-19 cases Ω and the observed data Ω^{ob} (a) and deaths Π and the observed data Π^{ob} (b) for high (dashed line) and low (continuous line) virulence sets. The estimated curve of Ω separates from the observed trend Ω^{ob} around on July 1 for low and on May 7 for high virulent variants; and the estimated curve of Π separates from the observed trend Π^{ob} around on July 21 for low and on May 16 for high virulent variants. The curves of Ω and Π for low virulent variant accompanied the trend of observed data in more 55 and 66 days than for higher one. It is worth stressing that, after splitting, the estimated curves of Ω and Π situate below the observed Ω^{ob} and Π^{ob} data for the lower virulent variant, but above the observed data for the higher virulent one.

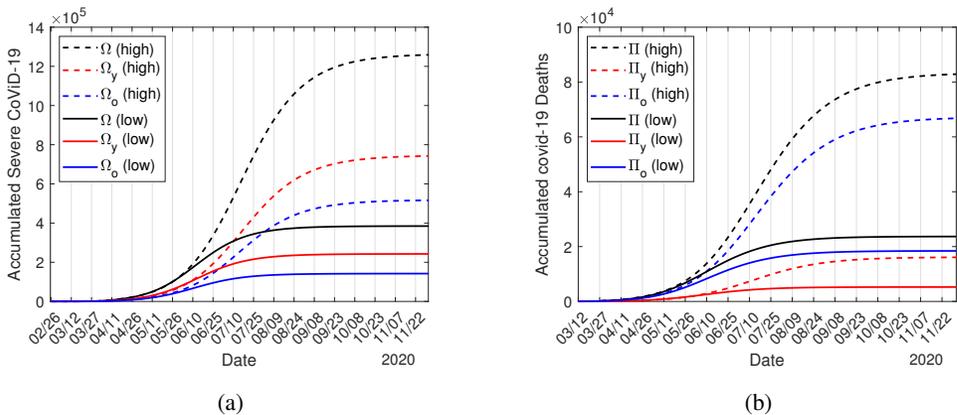


Figure 3: The extended curves until November 30, 2020: the accumulated Covid-19 cases Ω the observed data Ω^{ob} (a) and deaths Π the observed data Π^{ob} (b) for the higher (dashed line) and lower (continuous line) virulence. The curves correspond to the prediction of epidemic under quarantine and protective measures.

To assess the contributions of young and senior subpopulations in the overall epidemic, Figure 4 shows the extension of Figure 2 until November 30, 2020: accumulated Covid-19 cases Ω_y , Ω_o , and Ω (a), and the accumulated deaths Π_y , Π_o , Π (b) for high (dashed line) and low (continuous line) virulence sets. For low virulence (high virulence), the numbers of severe Covid-19 cases on May 13 are 242544 (742418) and 142104 (516293) for young and senior subpopulations, with the ratio being 1.56 (1.32). For low virulence (high virulence), the numbers of Covid-19 fatality cases on May 28 are 5254 (16080) and 18418 (66767) for young and senior subpopulations, with the ratio being 0.29 (0.24).

Figures 1 to 4 were obtained using $R_0 = 9.24$ for low and $R_0 = 7.02$ for high virulent variants, which illustrates the low transmissibility of high virulent one. Also, the number of deaths in high virulent variant epidemic is very huge compared to the low virulent one. Let us evaluate the ratio between asymptomatic and symptomatic individuals disregarding the age.

Figure 5 shows the curves of asymptomatic (A) and pre-symptomatic (P) individuals (a), and the estimated curves of mild (M) and severe (D) Covid-19 cases (b) for the high (dashed line) and

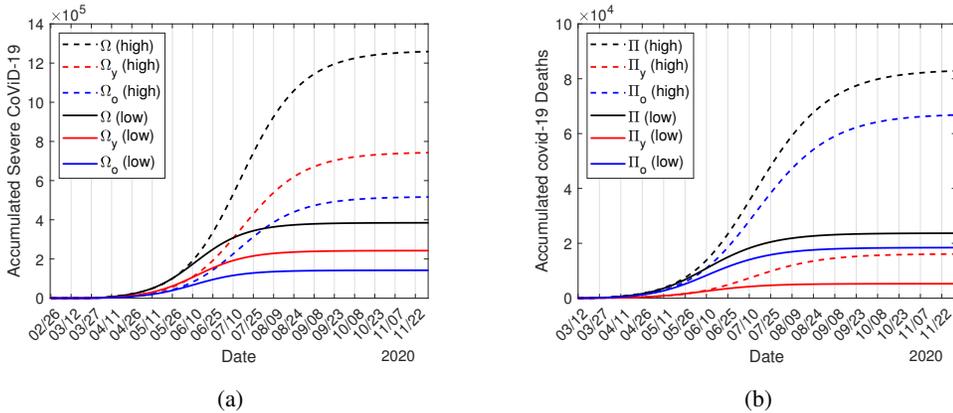


Figure 4: The extended curves until November 30, 2020: the accumulated Covid-19 cases Ω_y , Ω_o , and Ω (a), and the accumulated deaths Π_y , Π_o , Π (b) for the higher (dashed line) and lower (continuous line) virulence. The curves correspond to the prediction of epidemic under quarantine and protective measures.

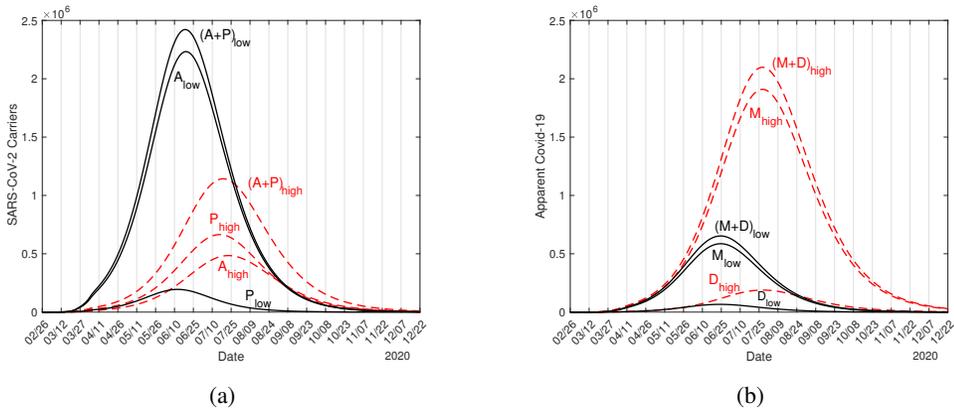


Figure 5: Estimated curves of asymptomatic (A) and pre-symptomatic (P) individuals (a), and the estimated curves of mild (M) and severe (D) Covid-19 cases (b) for the higher (dashed line) and lower (continuous line) virulence. The sum $A + P$ and $M + D$ are shown.

low (continuous line) virulence. Figure 5 also shows the curve of carriers capable of transmitting the virus ($A + P$) and the curve of apparent Covid-19 ($M + D$).

For the low virulence virus, the peaks of asymptomatic (A), pre-symptomatic (P), mild (M), and severe (D) Covid-19 are 2230000, 195000, 586000, and 66000, and for the high virulence virus, the peaks are 485000, 664000, 1910000, and 189000. (In the low virulence scenario, the peak of A increased by 460%, and the peaks of P , M , and D were reduced by 340%, 326%, and 286% compared to the high virulence values.) Asymptomatic and pre-symptomatic individuals

are carriers, which makes the control of infection more difficult. Comparing the lower and higher virulence, the ratio between the peaks of carriers ($A + P$) is 212%, and the ratio between peaks of apparent Covid-19 ($M + D$) is 31%.

Additionally, from the numerical simulations (Figure 5), the peaks of SARS-CoV-2 transmitters ($A + P + z_y M_y + z_o M_o$) for the low and high virulence scenarios are 2680000 and 1990000, occurring on June 19 and July 21. The ratio of low and high virulence peaks for all infectious individuals is 135%. The peaks of exposed (E), releasing virus ($A + P + M + D$), and harboring virus ($E + A + P + M + D$) individuals in the low virulence set were 1430000, 3060000, and 4410000, respectively, and for the high virulent variant, the peaks were 1210000 (85%), 3210000 (105%), and 4360000 (99%), respectively. The percentage shown in parentheses is the ratio high to low virulence. The number of individuals releasing the virus are slightly higher in the high virulence set, but the number of individuals carrying the virus is nearly equal under low or high virulence transmission. Notice that exposed individuals (new infection cases) are higher in low virulence due to a higher R_0 .

4 DISCUSSION

During the first months of the Covid-19 outbreak (input data set), the absence of mass testing of SARS-CoV-2 infection and clinical follow-up of infected individuals precluded the detection of asymptomatic and pre-symptomatic individuals. Considering this initial period, we explained São Paulo's Covid-19 data using two broadly different values representing the ratio between asymptomatic and symptomatic individuals (4 : 1 and 1 : 4). For each proportion l , we estimated the transmission rates β_y and β_o , and calculated the basic reproduction number R_0 . It is important to remember that the low and high virulent SARS-CoV-2 variants provided quite similar fitted Ω and Π curves (see Figure 1).

Figure 1 showed the parameters estimation. For high virulent variant, the reproduction number was $R_0 = 7.02$, with the partial numbers $R_{0y} = 6.14$ and $R_{0o} = 0.88$ ($R_{0o}/R_{0y} = 0.143$); while for low virulent variant, the reproduction number was $R_0 = 9.24$, with the partial numbers $R_{0y} = 7.73$ and $R_{0o} = 1.51$ ($R_{0o}/R_{0y} = 0.195$). The ratio between senior and young subpopulations is $N_{0o}/N_{0y} = 0.180$. The quotients between low and high virulent variants for R_0 , R_{0y} , and R_{0o} are 132%, 126%, and 171%, respectively. In an high virulent variant epidemic, the senior subpopulation is more protected when in interaction with young subpopulation than in a low variant epidemic. The effect of age on the epidemic was shown in Figures 2 and 4.

Remember that Figure 1 was estimated considering quarantine, and Figure 3 was extension of the estimated curve until November 30, 2020. In a direct observation, the predicted curve of severe Covid-19 cases for low variant epidemic matched the observed data 55 more days than the high virulent one (for fatality cases, more 66 days).

Let us discuss the model predictive ability by comparing the extended curves in Figure 3 with the observed data for the low and high virulence scenarios. However, it can be observed in Figure 3 that the low virulence epidemic fits very well into the subset of test data from May 13 to

June 30 (Ω) or July 20 (Π), which is the last day of the quarantine. The first observed point detaching from the low virulence curve Π occurs on July 21, while the severe Covid-19 curve Ω splits on July 1, 20 days earlier. Hence, this finding corroborates our assumption $\Delta = 15$ days in $\Pi(t + \Delta) = \alpha D(t)$ to estimate the fatality rates α_y and α_o . The effects of the relaxation appeared on July 1 and since then the observed number of deaths was systematically above the estimated curve of Π considering the quarantine only (relaxation began by mid-June 2020, and its effect on the epidemic curve D appeared 9 days later). Hence, the subset of test data from July 1 must be described by a model incorporating intermittent pulses of relaxation [23], which is not dealt with here. In [37], the relaxation incorporated in the model fitted very well the test data set considering the low virulent variant. We concluded that this epidemiological scenario can be explained by the spread of less virulent SARS-CoV-2 variant during the quarantine and relaxation (test data set).

Figure 5 showed enhanced virulence increasing the deaths due to Covid-19 by around 300%; however, the transmissibility measured by R_0 was reduced to $\sim 76\%$. The higher virulence increased by around 300% the number of mild and severe Covid-19 cases, potentially more transmissible but in isolation; however, the number of asymptomatic, pre-symptomatic, and a fraction of mild Covid-19 infected individuals transmitting the virus is reduced by 74%. The rate of SARS-CoV-2 transmission in the population is reduced, and the number of deaths goes up as virulence increases. Initially, the model variables (compartments) have similar values, but the peaks of the more virulent SARS-CoV-2 occur around one month later. In particular, the number of deaths in the more virulent epidemic is three times higher, although both variants presented similar fatality rates in the first three months.

However, during the test data set, besides the relaxation, the emerging of variants of SARS-CoV-2 may occur. The persistent transmission of SARS-CoV-2 should result in a plethora of new mutations and variants. Some of these variants acquired enhanced ability to infect cells and/or higher levels of replication, which may in turn increase virulence and, consequently, the risk of death. Moreover, as more viruses are released in the environment, the level of transmission also increases. However, individuals with severe symptoms are isolated in hospitals, and mild cases are tested and detected and advised to self-isolate. Conversely, asymptomatic and pre-symptomatic individuals (carriers) circulate freely and may get in close contact with susceptible individuals, unlike individuals manifesting any suspicious symptoms (not necessarily Covid-19). This fact was exposed by our model, that showed that the presence of high virulence viruses results in more deaths but is less transmissible, i.e., display a lower R_0 . Therefore, while the epidemic curve increases slowly, the curve of deaths increases much more quickly than in the low virulence scenario.

Indeed, the gamma variant was first detected in Manaus by November 2020. At the end of February 2021, São Paulo still maintained part of the population isolated, but the accumulated number of deaths of 68904 on March 24, 2021, surpassed the 57300 predicted by the epidemic with a low virulence SARS-CoV-2 variant. This high number of deaths is likely a consequence of the appearance of more virulent variants (but not so high as $l = 0.2$ as shown in Figures 1 to 5) during this period [10]. Based on the model's analysis, we may assume that more virulent viruses were

present in the pool of SARS-CoV-2 variants, resulting in more severe Covid-19 cases and deaths. It should be stressed out, however, that an increased number of severe cases and deaths caused by the highly virulent variant does not imply increased transmissibility, according to our model. The omicron variant appeared in South Africa in November 2021. Nevertheless, the omicron variant was showed to be more transmissible, but having less virulence [35].

5 CONCLUSION

In conclusion, throughout the course of an epidemic fast-mutating RNA viruses acquire new mutations resulting in a pool of variants. The new variants may display lower or higher levels of virulence than the original SARS-CoV-2 strain. In a short-run epidemic, the spread of highly virulent variants increase the number of severe cases and deaths, but, paradoxically, reduce viral transmission. However, in the long-run highly-transmissible low virulence variants will prevail. The competition between two strains is left to further work (see [31] for the drug-sensitive and resistant *M. tuberculosis* transmission modeling). Therefore, for the short-run epidemic control, strict isolation of symptomatic (and perhaps also of suspected cases) should be adopted to avoid the spread of highly virulent SARS-CoV-2 variants which will otherwise result in more severe cases and deaths.

REFERENCES

- [1] Greenhalgh T, *et al.* Ten scientific reasons in support of airborne transmission of SARS-CoV-2. *Lancet* 2021; 397 (10285); 1603-1605.
- [2] Yuan L, *et al.* Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. Published April 27, 2020. Available at: <https://doi.org/10.1038/s41586-020-2271-3>(accessed May 10, 2020).
- [3] Zhu N, *et al.* A novel Coronavirus from patients with pneumonia in China, 2019. *N. Eng. J. Med.* 2020; 381: 1-7. doi: 10.1056/NEJMoa2001017.
- [4] WHO, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (Covid-19), 16-24 February 2020 (2020).
- [5] Sender R, *et al.* The total number and mass of SARS-CoV-2 virions. *PNAS* 2021; 118 (25) e2024815118. doi: 10.1073/pnas.2024815118.
- [6] WHO, Coronavirus Disease (Covid-19): Virus Evolution. 30 December 2020. Available at: <https://www.who.int/news-room/q-a-detail/sars-cov-2-evolution> (accessed February 23, 2021).
- [7] CDC, SARS-CoV-2 Variants. Updated Jan. 31, 2021. Available at: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/variant-surveillance/variant-info.html> (accessed February 23, 2021).

- [8] Korber B, *et al.* Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the Covid-19 virus. *Cell* 2021; 182: 812-817. doi: <https://doi.org/10.1016/j.cell.2020.06.043>.
- [9] Davies NG, *et al.* Estimated transmissibility and severity of novel SARS-CoV-2 Variant of Concern 202012/01 in England. *MedRxiv* 2020. doi:10.1101/2020.12.24.20248822.
- [10] Faria NR, *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* 2021; 372 (6544); 815-821.
- [11] Martin DP, *et al.* The emergence and ongoing convergent evolution of the N501Y lineages coincides with a major global shift in the SARS-CoV-2 selective landscape. *MedRxiv* 2021. doi: 10.1101/2021.02.23.21252268.
- [12] Naveca FG, *et al.* COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nature Medicine* 2021; 27; 1230-1238.
- [13] Zhan Y, Yin H, Yin JY. B.1.617.2 (Delta) Variant of SARS-CoV-2: features, transmission and potential strategies. *Int J Biol Sci* 2022;18(5); 1844-1851.
- [14] Planas D, *et al.* Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* 2021; 596; 276-280.
- [15] Jansen L, *et al.* Investigation of a SARS-CoV-2 B.1.1.529 (Omicron) Variant Cluster — Nebraska, November–December 2021. *Morbidity and Mortality Weekly Report* 2021; 70 (5152);1782-1784.
- [16] Kannan SR, *et al.* Omicron SARS-CoV-2 variant: Unique features and their impact on pre-existing antibodies. *Journal of Autoimmunity* 2022; 126; 102779. doi: 10.1016/j.jaut.2021.102779.
- [17] Saxena SK, *et al.* Characterization of the novel SARS-CoV-2 Omicron (B.1.1.529) variant of concern and its global perspective. *J Medical Virology* 2021. <http://doi.org/10.1002/jmv.27524>.
- [18] Kumar S, Thambiraja TS, Karuppanan K, Subramaniam G. Omicron and Delta variant of SARS-CoV-2: A comparative computational study of spike protein. *J Medical Virology* 2021. <http://doi.org/10.1002/jmv.27526>.
- [19] Yang HM, Lombardi Junior LP, Campos AC. Are the SIR and SEIR models suitable to estimate the basic reproduction number for the Covid-19 epidemic? *MedRxiv* 2020. doi: <https://doi.org/10.1101/2020.10.11.20210831>.
- [20] Arons MM, *et al.* Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *The New Engl. Jour. Medicine* 2020; April 24, 2020. doi: 10.1056/NEJMoa2008457.

- [21] Anderson RM, May, RM. Infectious Diseases of Human. Dynamics and Control. Oxford, New York, Tokyo: Oxford University Press; 1991: 757 p.
- [22] Yang HM, Lombardi Junior LP, Castro FFM, Yang AC. Mathematical modeling of the transmission of SARS-CoV-2 – Evaluating the impact of isolation in São Paulo State (Brazil) and lockdown in Spain associated with protective measures on the epidemic of Covid-19. PLoS ONE 2021; 16(6): e0252271. <https://doi.org/10.1371/journal.pone.0252271>.
- [23] Yang HM, Lombardi Junior LP, Castro FFM, Yang AC. Mathematical model describing Covid-19 in São Paulo State, Brazil – Evaluating isolation as control mechanism and forecasting epidemiological scenarios of release. Epidemiology and Infection 2020 148: e155. doi: 10.1017/S0950268820001600.
- [24] SEADE, SP contra o novo coronavírus – Boletim completo. Available from: https://www.seade.gov.br/coronavirus/?utm_source=portal&utm_medium=banner&utm_campaign=boletim-completo (accessed August 14, 2020).
- [25] Alvarenga MY, Sameshima K, Baccalá LA, Yang HM. Non-linear analysis of the rhythmic activity in rodent brain. Mathematical Biosciences 1999; 157 (1-2): 287-302.
- [26] Informações Covid-19. <http://www.transparencia.am.gov.br/Covid-19/monitoramento-Covid-19/> (accessed February 27, 2021).
- [27] Freitas ARR, *et al.* The emergence of novel SARS-CoV-2 variant P.1 in Amazonas (Brazil) was temporally associated with a change in the age and sex profile of COVID-19 mortality: A population based ecological study. The Lancet Regional Health - Americas 2021; 2; 100064. <https://doi.org/10.1016/j.lana.2021.100021>.
- [28] Barbosa GR, *et al.* Rapid spread and high impact of the Variant of Concern P.1 in the largest city of Brazil. MedRxiv 2021. doi: 10.1101/2021.04.10.21255111.
- [29] Francisco Junior RS, *et al.* Turnover of SARS-CoV-2 Lineages Shaped the Pandemic and Enabled the Emergence of New Variants in the State of Rio de Janeiro, Brazil. Viruses 2021; 13 (10); 2013. doi: 10.3390/v13102013
- [30] Martins AF, *et al.* Detection of SARS-CoV-2 lineage P.1 in patients from a region with exponentially increasing hospitalisation rate, February 2021, Rio Grande do Sul, Southern Brazil. Eurosurveillance 2021; 26 (12). doi: 10.2807/1560-7917.es.2021.26.12.2100276
- [31] Raimundo SM, Yang HM, Venturino E. Theoretical assessment of the relative incidence of sensitive and resistant tuberculosis epidemic in presence of drug treatment. Math. Biosc. Eng. 2014; 11(4): 971-993.
- [32] Diekmann O, Heesterbeek JAP, Roberts MG. The construction of next-generation matrices for compartmental epidemic models. J. R. Soc. Interface 2010; 7: 873-885.

- [33] Yang HM. The basic reproduction number obtained from Jacobian and next generation matrices – A case study of dengue transmission modelling. *BioSystems* 2014; 126: 52-75.
- [34] Yang HM, Greenhalgh D. Proof of conjecture in: The basic reproduction number obtained from Jacobian and next generation matrices – A case study of dengue transmission modelling. *Appl. Math. Comput.* 2015; 265: 103-107.
- [35] Bálint G, Vörös-Horváth B and Széchenyi A. (2022). Omicron: increased transmissibility and decreased pathogenicity. *Signal Transduction and Targeted Therapy* 2022; 7(1): 1-3.
- [36] Fundação Sistema Estadual (SEADE) database (<https://www.seade.gov.br>) (accessed 9 April 2020).
- [37] Yang HM, Lombardi Junior LP, Castro FFM, Campos AC. Modeling the transmission of the new coronavirus in São Paulo State, Brazil – assessing the epidemiological impacts of isolating young and elder persons. *Mathematical Medicine and Biology: A Journal of the IMA* 2021; 38 (2): 137-177. doi: <https://doi.org/10.1093/imammb/dqaa015>.
- [38] Yang HM, Lombardi Junior LP, Castro FFM, Yang AC. Evaluating the impacts of relaxation and mutation in the SARS-CoV-2 on the COVID-19 epidemic based on a mathematical model: a case study of São Paulo State (Brazil). *Comput. Appl. Math.* 2021; 40(8): 272 doi: 10.1007/s40314-021-01661-w.



APPENDIX A SARS-COV-2 AND VARIANTS OF CONCERN

Like other RNA-based viruses, SARS-CoV-2 mutation rate is relatively high (about 3×10^{-6} mutations per nucleotide per replication cycle [5]) but not as high as in other RNA viruses such as HIV or influenza [6]. “Viruses of a particular lineage that acquire one or more non-synonymous new mutations are referred to as variant of the original virus” [6]. Some of these new variants are of public health concern (known as ‘variants of concern’ or VOC) either because they have an increased capacity of causing severe illness or because they are more efficiently transmitted than other variants. For instance, the SARS-CoV-2 VOCs alpha (B.1.1.7), gamma (P.1 or B.1.1.28), delta (B.1.617) and omicron, that emerged in the United Kingdom, Brazil, India and South Africa, respectively (see [7] and references therein). Some of these VOCs display high lethality, high levels of transmission and are able to escape the host immune system. Indeed, in many cases more virulent variants of SARS-CoV-2 are also more transmissible [8].

Given the relatively high SARS-CoV-2 mutation rate the emergence of new variants is expected in the course of a pandemic. However, only a few of them are expected to confer significant selective advantage to the virus. Some of these new variants are of public health concern either because they have an increased capacity of causing severe illness or because they are more efficiently transmitted than other variants.

The original SARS-CoV-2 variant was first detected in Wuhan, China by December 2019. The first VOC - alpha, appeared in the United Kingdom less than one year later, and quickly spread throughout the country. The alpha variant carries 23 mutations, including 14 non-synonymous substitutions, 3 deletions, and 6 synonymous mutations, three of them targeting the spike protein. These mutations are related to a 50% increase in transmissibility (according to the CDC), making the virus highly contagious and difficult to contain [9]. The gamma VOC was first detected in Manaus by November 2020 [10]. From there, this variant disseminated throughout the country creating havoc and a pattern of high lethality. The gamma variant harbours a total of 25 mutations (synonymous and non-synonymous) compared with its most closely related ancestral, including 10 amino acid substitutions in the spike protein, some of them associated with a higher degree of transmissibility [10] [11]. In addition, the viral load in patients infected with the gamma variant was ~ 10 -fold higher than in those with other SARS-CoV-2 variants [12]. In Brazil alone, the gamma variant was responsible for the death of about 400,000 people, in a 6-month period. The delta variant carries 4 mutations of interest in the spike protein. It was first detected in India in late 2020. This VOC also displayed increased virulence and transmissibility [13], but its mutations have also enabled the virus to evade immune response, and for that reason most vaccines were less effective against this variant [14]. Finally, the omicron variant, that appeared in South Africa in November 2021, is the most phylogenetically divergent VOC to date. It is associated with enhanced transmissibility, reduced vaccine-induced immunity and increased risk of re-infection. Omicron harbors more than 50 mutations, 34 of them in the spike protein: 30 amino acid substitutions, 3 deletions, and 1 small insertion [15]. Half of these amino acid substitutions occur in the receptor binding domain, which might explain the high transmissibility of the this variant [16, 17, 18].

APPENDIX B THE STEADY-STATE ANALYSIS OF THE MODIFIED SQEAPMDR MODEL

The system of equations (2.3), (2.4) and (2.5), time-varying population, approach an asymptote, which is not an equilibrium value. However, in term of fractions this system attains steady-state. Defining the fraction $x_j = X_j/N$, for $j = y, o$, with $X_j = \{S_j, Q_j, E_j, A_j, P_j, M_j, D_j, R\}$, we have, using equation (2.6),

$$\frac{d X_j}{dt N} = \frac{1}{N} \frac{d}{dt} X_j - \frac{X_j}{N} \frac{1}{N} \frac{d}{dt} N = \frac{1}{N} \frac{d}{dt} X_j - x_j (\phi - \mu - \alpha_y d_y - \alpha_o d_o),$$

and the system of equations (2.3), (2.4) and (2.5) in terms of fractions become, for susceptible individuals,

$$\begin{cases} \frac{d}{dt} s_y &= \phi - (\varphi + \phi) s_y - \lambda s_y + s_y (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} s_o &= \varphi s_y - \phi s_o - \lambda \psi s_o + s_o (\alpha_y d_y + \alpha_o d_o), \end{cases} \tag{B.1}$$

for isolated and infected individuals,

$$\begin{cases} \frac{d}{dt} q_j &= -\phi q_j + q_j (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} e_j &= \lambda (\delta_{jy} + \psi \delta_{jo}) s_j - (\sigma_j + \phi) e_j + e_j (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} a_j &= l_j \sigma_j e_j - (\gamma_j + \phi) a_j + a_j (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} p_j &= (1 - l_j) \sigma_j e_j - (\gamma_{1j} + \phi) p_j + p_j (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} m_j &= (1 - \chi_j) \gamma_j a_j + k_j \gamma_{1j} p_j - (\gamma_{3j} + \phi) m_j + m_j (\alpha_y d_y + \alpha_o d_o) \\ \frac{d}{dt} d_j &= (1 - k_j) \gamma_{1j} p_j - (\gamma_{2j} + \phi + \alpha_j) d_j + d_j (\alpha_y d_y + \alpha_o d_o), \end{cases} \tag{B.2}$$

and for recovered individuals,

$$\frac{d}{dt} r = \chi_y \gamma_y a_y + \gamma_{3y} m_y + \gamma_{2y} d_y + \chi_o \gamma_o a_o + \gamma_{3o} m_o + \gamma_{2o} d_o - \phi r + r (\alpha_y d_y + \alpha_o d_o), \tag{B.3}$$

where λ is the force of infection given by equation (2.2) re-written as

$$\lambda = \beta_{1y} a_y + \beta_{2y} p_y + \beta_{3y} z_y m_y + \beta_{1o} a_o + \beta_{2o} p_o + \beta_{3o} z_o m_o,$$

letting $\varepsilon = 1$ (absence of protective measures), with

$$\sum_{j=y,o} (s_j + q_j + e_j + a_j + p_j + m_j + d_j) + r = 1.$$

This new system of equations has steady-states (the sum of derivatives of all classes is zero).

The trivial (disease-free) equilibrium point P^0 of the new system of equations (B.1), (B.2) and (B.3) is given by

$$P^0 = (s_j^0, q_j^0 = 0, e_j^0 = 0, a_j^0 = 0, p_j^0 = 0, m_j^0 = 0, d_j^0 = 0, r^0 = 0),$$

for $j = y$ and o , where

$$\begin{cases} s_y^0 = \frac{\phi}{\phi + \varphi} \\ s_o^0 = \frac{\varphi}{\phi + \varphi}, \end{cases} \tag{B.4}$$

with $s_y^0 + s_o^0 = 1$.

Let us assess the stability of P^0 by applying the next generation matrix theory considering the vector of variables $x = (e_y, a_y, p_y, m_y, e_o, a_o, p_o, m_o)$ [32]. We apply method proposed in [33] and proved in [34]. To obtain the basic reproduction number, diagonal matrix V is considered. Hence, the vectors f and v are

$$f^T = \begin{pmatrix} \lambda s_y + e_y (\alpha_y d_y + \alpha_o d_o) \\ l_y \sigma_y e_y + a_y (\alpha_y d_y + \alpha_o d_o) \\ (1 - l_y) \sigma_y e_y + p_y (\alpha_y d_y + \alpha_o d_o) \\ (1 - \chi_y) \gamma_y a_y + k_y \gamma_{1y} p_y + m_y (\alpha_y d_y + \alpha_o d_o) \\ \lambda \psi s_o + e_o (\alpha_y d_y + \alpha_o d_o) \\ p_o \sigma_o e_o + a_o (\alpha_y d_y + \alpha_o d_o) \\ (1 - p_o) \sigma_o e_o + p_o (\alpha_y d_y + \alpha_o d_o) \\ (1 - \chi_o) \gamma_o a_o + k_o \gamma_{1o} p_o + m_o (\alpha_y d_y + \alpha_o d_o) \end{pmatrix} \quad \text{and} \quad v^T = \begin{pmatrix} (\sigma_y + \phi) e_y \\ (\gamma_y + \phi) a_y \\ (\gamma_{1y} + \phi) p_y \\ (\gamma_{3y} + \phi) m_y \\ (\sigma_o + \phi) e_o \\ (\gamma_o + \phi) a_o \\ (\gamma_{1o} + \phi) p_o \\ (\gamma_{3o} + \phi) m_o \end{pmatrix},$$

where the superscript T stands for the transposition of a matrix, from which we obtain the matrices F and V (see [32]) evaluated at the trivial equilibrium P^0 , which were omitted. The next generation matrix FV^{-1} is

$$FV^{-1} = \begin{bmatrix} 0 & \frac{\beta_{1y}s_y^0}{\gamma_y+\phi} & \frac{\beta_{2y}s_y^0}{\gamma_{1y}+\phi} & \frac{\beta_{3y}z_y s_y^0}{\gamma_{3y}+\phi} & 0 & \frac{\beta_{1o}s_y^0}{\gamma_o+\phi} & \frac{\beta_{2o}s_y^0}{\gamma_{1o}+\phi} & \frac{\beta_{3o}z_o s_y^0}{\gamma_{3o}+\phi} \\ \frac{l_y \sigma_y}{\sigma_y+\phi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{(1-l_y) \sigma_y}{\sigma_y+\phi} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{(1-\chi_y) \gamma_y}{\gamma_y+\phi} & \frac{k_y \gamma_{1y}}{\gamma_{1y}+\phi} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\beta_{1y} \psi s_o^0}{\gamma_y+\phi} & \frac{\beta_{2y} \psi s_o^0}{\gamma_{1y}+\phi} & \frac{\beta_{3y} z_y \psi s_o^0}{\gamma_{3y}+\phi} & 0 & \frac{\beta_{1o} \psi s_o^0}{\gamma_o+\phi} & \frac{\beta_{2o} \psi s_o^0}{\gamma_{1o}+\phi} & \frac{\beta_{3o} z_o \psi s_o^0}{\gamma_{3o}+\phi} \\ 0 & 0 & 0 & 0 & \frac{p_o \sigma_o}{\sigma_o+\phi} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{(1-p_o) \sigma_o}{\sigma_o+\phi} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{(1-\chi_o) \gamma_o}{\gamma_o+\phi} & \frac{k_o \gamma_{1o}}{\gamma_{1o}+\phi} & 0 \end{bmatrix}$$

and the characteristic equation corresponding to FV^{-1} is

$$\varkappa^5 [\varkappa^3 - (R_{1y}s_y^0 + R_{1o}s_o^0) \varkappa - (R_{2y}s_y^0 + R_{2o}s_o^0)] = 0,$$

with the basic reproduction number R_0 being given by

$$R_0 = (R_{1y} + R_{2y})s_y^0 + (R_{1o} + R_{2o})s_o^0, \tag{B.5}$$

where the initial fractions s_y^0 and s_o^0 are given by equation (B.4), and the partial basic reproduction numbers R_{1y} , R_{2y} , R_{1o} , and R_{2o} are given by equation (2.11) in the main text. The spectral

radius $\rho(FV^{-1})$ is the biggest solution of a third-degree polynomial, not easy to evaluate. The procedure proposed in [33] allows us to obtain the threshold R_0 as the sum of coefficients of the characteristic equation. Hence, the trivial equilibrium point P^0 is locally asymptotically stable if $R_0 < 1$.