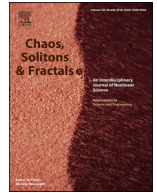




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations

Sebastián Contreras<sup>a,b,\*</sup>, H. Andrés Villavicencio<sup>b</sup>, David Medina-Ortiz<sup>b,c</sup>,  
Juan Pablo Biron-Lattes<sup>b,d</sup>, Álvaro Olivera-Nappa<sup>b,d,\*</sup>

<sup>a</sup> Laboratory for Rheology and Fluid Dynamics, Universidad de Chile, Beauchef 850, Santiago 8370448, Chile

<sup>b</sup> Centre for Biotechnology and Bioengineering, Universidad de Chile, Beauchef 851, Santiago 8370448, Chile

<sup>c</sup> Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>d</sup> Department of Chemical Engineering, Biotechnology and Materials, Universidad de Chile, Beauchef 851, Santiago 8370448, Chile

## ARTICLE INFO

### Article history:

Received 23 April 2020

Accepted 20 May 2020

Available online 25 May 2020

### Keywords:

COVID-19 pandemic

SARS-CoV2

Multigroup model

Public-health

SEIRA models

## ABSTRACT

The outbreak and propagation of COVID-19 have posed a considerable challenge to modern society. In particular, the different restrictive actions taken by governments to prevent the spread of the virus have changed the way humans interact and conceive interaction. Due to geographical, behavioral, or economic factors, different sub-groups among a population are more (or less) likely to interact, and thus to spread/acquire the virus. In this work, we present a general multi-group SEIRA model for representing the spread of COVID-19 among a heterogeneous population and test it in a numerical case of study. By highlighting its applicability and the ease with which its general formulation can be adapted to particular studies, we expect our model to lead us to a better understanding of the evolution of this pandemic and to better public-health policies to control it.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Since the outbreak of novel SARS-CoV2 in late 2019, the world's reaction capacity has been challenged in every possible way. From the first self-isolation measures to total lock-downs and quarantines for preventing its spreading, the way humans interact has changed, and probably, for good. Joined efforts from the scientific community have helped to design strategies for supporting public-health policies, and for providing a better understanding of the current scenario. In particular, several mathematical models have been proposed in recent weeks to fit public databases on the SARS-CoV2 outbreak and forecast its evolution. Despite their particularities, most of them evolve from the well-known SIR model proposed by Kermack and McKendrick [12]. The basic idea behind this model, and its variants, is to divide a population in different compartments, representing the  $S$  (susceptible),  $E$  (exposed),  $I$  (infected),  $A$  (asymptomatic),  $R$  (recovered),  $D$  (dead),  $Q$  (quarantined), among other fractions of it. Centering our analysis on the spread of COVID-19, and restricting our search to those models recently proposed, we may find different combinations of the variables mentioned above.

Even though research has been done on SIR and SIRD models for COVID-19, with different levels of complexity— as in [4,6,20], among others, for SIR, and in [3,5,9], among others, for SIRD—, properties of the virus and the way modern society interacts are more likely to be appropriately represented by SEIR/SEIRA models.

Yang et al. [23] presented a modified SEIR model, where variables  $S$  and  $E$  had an in-out function, accounting for quarantine, and  $E$  was considered to be infectious. Even though the model involved several variables, the parameter-fitting stage only considered an exponential approximation, leaving unclear how the infection and recovery rates were decoupled. Another variant of SEIR models was presented by Yang and Wang [22], including an equation for the concentration of SARS-CoV2 in the environmental reservoir. [19] present a sophisticated 7-compartment variant of SEIR models, where the extra variables account for the quarantined  $Q$ , dead  $D$ , and insusceptible  $P$  fractions of the population. Moreover, their model includes the temporal evolution of the rates of recovery and infection-driven death. Kucharski et al. [13] used a stochastic dynamic model to fit publicly available datasets of Wuhan to give better insights on the early dynamics of COVID-19, and therefore predict if newly introduced cases would generate outbreaks in other areas. Pang [18] present several public health recommendations, based on the modeling of the effect of different regulations for preventing the spread of COVID-19 on a SEIRA model similar to the one presented by Aguilar et al. [1],

\* Corresponding authors.

E-mail addresses: [scontreras@cebib.cl](mailto:scontreras@cebib.cl) (S. Contreras), [aolivera@ing.uchile.cl](mailto:aolivera@ing.uchile.cl) (Á. Olivera-Nappa).

highlighting the challenges that an epidemic of an aerial contagious virus with asymptomatic patients pose.

Nevertheless, the different models listed above do not account for the several factors that could result in a heterogeneous spread of the infection in the population. Barriers can be risen by the geospatial configuration of a country and differential connectivity between its regions, or even by the way the different communities or social classes interact in a city. Even though work has been done in multi-group SIR models [11] and the stability of their endemic equilibrium [10,21], the particularities of the COVID-19 pandemics and human interaction networks require to be modeled to be correctly represented.

In this work, we present a general multi-group SEIRA model for representing the spread of novel COVID-19 through populations with heterogeneous characteristics, such as a heavily centralized organization with poor connections between the provinces, substantial social inequality among the population, or age/behavioral groups. Due to the generality of our model, we can easily modify it to account for different particularities of any population. In our model, we represent the interactions between the different groups by a non-necessarily symmetrical matrix connecting the different sub-groups, accounting for the fraction that effectively interacts. Particularities of novel COVID-19 were used to model the infection dynamics, so individuals could be infected by foreigners in their region, or while they are visiting a different one. We apply our model in a case of study for representing the dynamics between different behavioral groups, highlighting its applicability to the early prediction of contingencies.

### 2. Our approach

The foundations on which our model relies are the well-known mathematical and biological hypotheses for SEIR single and multi-group models [11]. Other suppositions, as listed below, are based on the state-of-the-art knowledge of COVID-19 and other population-balance considerations.

- There exists an incubation period for the virus, which average has been reported to range between 5.2 days [14,15] and 6.4 days [2]. This fact justifies the existence of a compartment  $E$  (exposed).
- The asymptomatic fraction of a given population, which varies from 29% [16], going through 50% [17] and rising between 50 and 75% [7], will be considered constant  $\alpha$ . This hypothesis makes sense if the size of the population and its immunological characteristics are homogeneous.
- The only possible interaction among different classes is infection. Recoveries and deaths are supposed to occur in the group where they belong.
- Asymptomatic patients are  $\xi$  times more likely to transmit the infection successfully to a healthy patient. Even though the viral load of asymptomatic individuals is in the same order of magnitude than in symptomatic patients [24], our supposition relies on a social-interaction basis. Symptomatic patients are easily avoided, quarantined, or likely to choose self-isolation, which is not the case of asymptomatic individuals.

Our model also accounts for Asymptomatic individuals  $A$ , but as we suppose them to represent a constant  $\alpha$  fraction of the total infected individuals, they are considered in the equation for  $I$ .

$$I_{total} = \underbrace{(1 - \alpha)I}_{\text{Symptomatic}} + \underbrace{\alpha I}_{\text{A: Asymptomatic}} \quad (1)$$

The general idea behind our model, and workflow of its network-scheme, is presented in Fig. 1. The particular deduction

and novelties involved in the SEIR equations are presented in the following section.

### 3. The model

As the different sub-groups interact, we expect to find in the  $i$ 'th group representatives of the other groups. As each group is subdivided into classes according to the SEIRA model, the total amount of a given  $X$  class,  $X = \{S, E, I, R, n\}$ , present in a group  $i$  is given by :

$$X_i^T = \underbrace{X_i \left( 1 - \sum_{j \neq i} \Phi_X^{ij} \right)}_{\text{Locals staying in their own district}} + \underbrace{\sum_{j \neq i} \Phi_X^{ji} X_j}_{\text{visitors (floating population)}} \quad (2)$$

where  $\Phi^{ji}$  represents the fraction of the  $j$ 'th class present in class  $i$ . As the maximum amount of individuals that may leave their class is the total of it, the following restriction must be met:

$$\sum_{j \neq i} \Phi^{ij} \leq 1. \quad (3)$$

By defining an "auto-fraction"  $\Phi_X^{ii}$ , the notation is considerably enlightened, as a single term would account for the whole population.

$$\Phi_X^{ii} = 1 - \sum_{k \neq i} \Phi_X^{ik} \Rightarrow X_i^T = \sum_{j=1}^n \Phi_X^{ji} X_j. \quad (4)$$

For studying variations of the susceptible class among the different populations, we must consider the factors which modify it:

$$\frac{dS_i}{dt} = \text{Natural net growth} - \text{Intraclass infections} - \text{Interclass infections}, \quad (5)$$

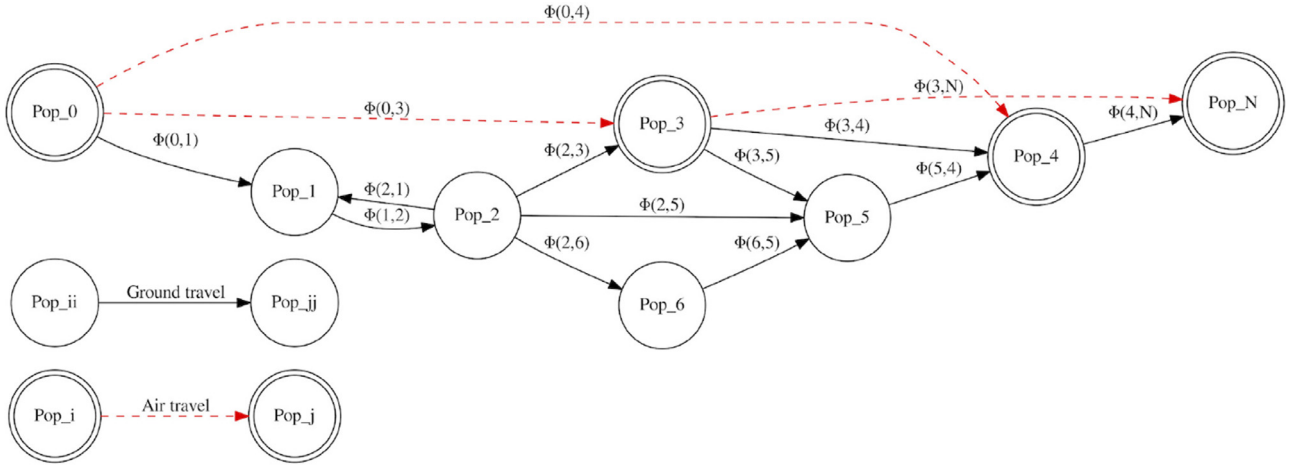
Only with the aim of providing a general framework, easily extendable to other situations, we can also consider the natural birth-death dynamics of the population, characterized by  $\Lambda, p_i$ , and  $d_i$ . Note that we can neglect those terms if the dynamics of the infection are faster than them. The interesting part is in the dynamics behind contagion, as schematized in Fig. 2.

Consider the fraction  $\Phi^{ij} n_i$ , representing the part of the  $i$ 'th population that is among  $j$ . There, this group would interact with the total population present at a time (locals and floating population), among whom infected people could be found (red individuals in Fig. 2) from all the different classes (the color of their aura). The contagion process can be modeled by Eq. (6)

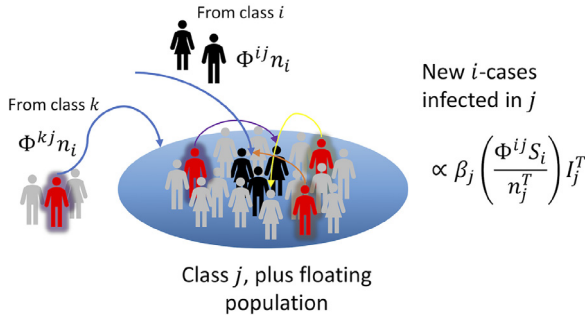
$$\text{New infections} \propto \underbrace{\beta_i}_{\text{Local infection rate}} \cdot \underbrace{\left( \frac{\Phi_S^{ij} S_i}{n_j^T} \right)}_{\text{Density of } i\text{-individuals in } j} \cdot \underbrace{\left( \sum_{k=1}^n \Phi_I^{kj} I_k \right)}_{\text{Total infected individuals in } j} \quad (6)$$

Note that Eq. (6) also accounts for the intra-class contagion, when defining  $\Phi_i^{ii}$  in accordance to Eq. (4). In particular, supposing new infections passed by asymptomatic individuals are  $\xi$  times more likely to occur, and all the other variables following a typical SEIR scheme, we may write the total dynamics for the  $i$ 'th group:

$$\frac{dS_i}{dt} = (1 - p_i) \Lambda_i - d_i S_i - (1 + (\xi - 1) \alpha) \sum_{j=1}^n \beta_j \left( \frac{\Phi_S^{ij} S_i}{n_j^T} \right) \left( \sum_{k=1}^n \Phi_I^{kj} I_k \right) \quad (7)$$



**Fig. 1.** Our approach proposes the partition of a heterogeneous population into several (as-many-as-required) sub-populations, where the hypotheses for SEIR models are satisfied. The different populations share common characteristics, as a geographical zone (this scheme), but not restricted only to that interpretation. This same reasoning (and model) can be applied to different behavioral groups, social classes, and age groups, through an appropriate interpretation for the interaction function.



**Fig. 2.** Schematic representation of the contagion processes between classes. According to the form of the interaction matrix  $\Phi$ , individuals from class  $i$  would contribute to the floating population of class  $j$ , where they would interact with individuals from -in principle- all other classes. As among them might be infected individuals from all origins, the interaction term should be corrected as presented.

$$\frac{dE_i}{dt} = (1 + (\xi - 1)\alpha) \sum_{j=1}^n \beta_j \left( \frac{\Phi^{ij}S_i}{n_j^T} \right) \left( \sum_{k=1}^n \Phi^{kj}I_k \right) - (\epsilon_i + d_i)E_i \tag{8}$$

$$\frac{dI_i}{dt} = \epsilon_i E_i - (\gamma_i + \theta_i + d_i)I_i \tag{9}$$

$$\frac{dR_i}{dt} = \gamma_i I_i - d_i R_i \tag{10}$$

Note that Eq. (9) is the result of considering the dynamics of both symptomatic  $((1 - \alpha)I)$  and asymptomatic  $(\alpha I)$  individuals together.

An important remark on the nature of the interactive term of Eq. (7), is that the different involved variables are fully coupled:

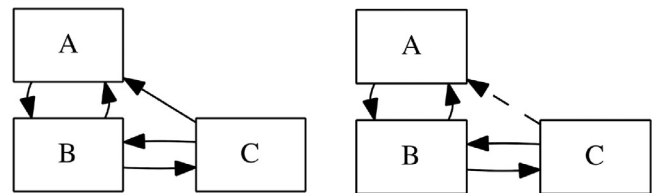
$$\sum_{j=1}^n (1 + (\xi - 1)\alpha)\beta_j \left( \frac{\Phi^{ij}S_i}{n_j^T} \right) \left( \sum_{k=1}^n \Phi^{kj}I_k \right) = \sum_{k=1}^n \beta_k f(S_i, \mathbf{X}, \Phi). \tag{11}$$

In the special case of not considering births and deaths from other causes and if all the different fluxes  $\Phi^{ij}$  are zero (for  $j \neq i$ ), function  $f$  would be linear for  $I_j$ . In any other situation,  $f(S_i, \mathbf{X}, \Phi)$  would be a non linear function not only of  $S_i$  and  $I_k$ , but of all the involved variables  $X_j$ . Nevertheless, if the flux structure of the network  $\Phi$  remains constant (or it is constant by parts), function  $f$  satisfies the hypothesis presented in Sun and Shi [21] for the global stability of the endemic equilibrium.

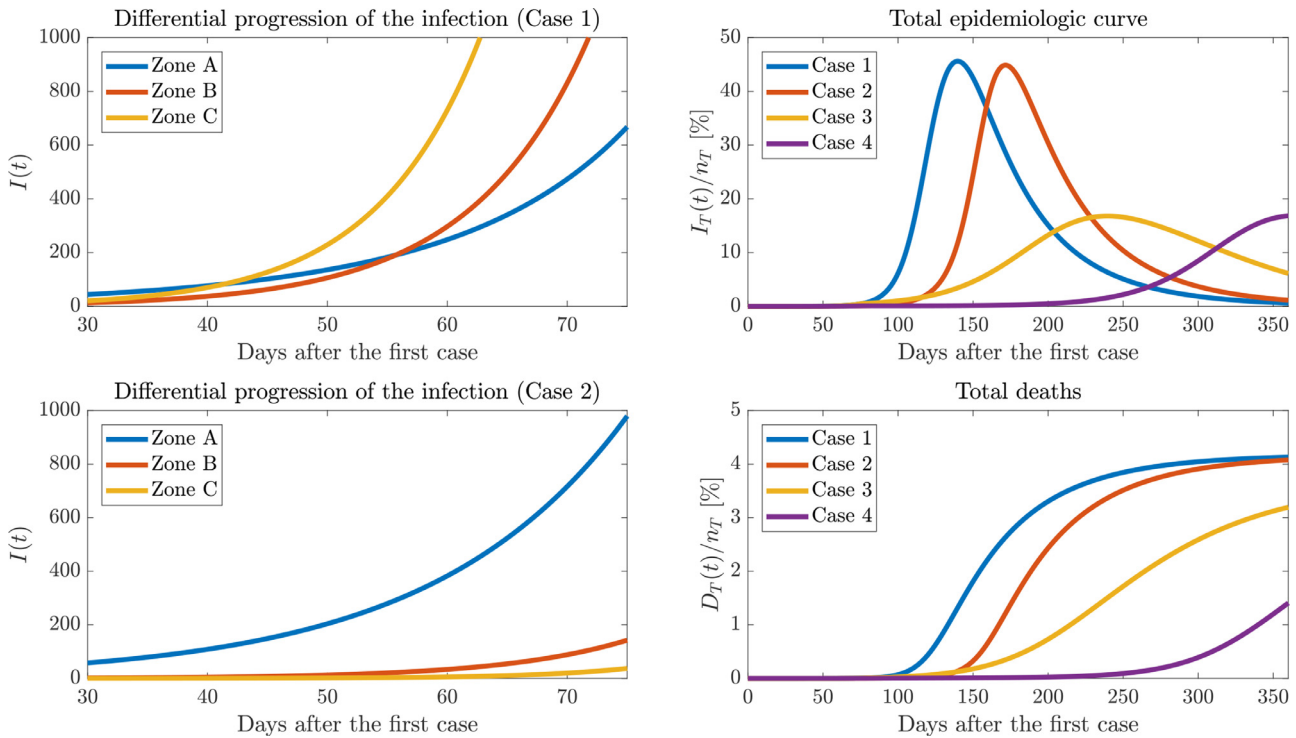
#### 4. Case of study: Mutual dependency between heterogeneous groups

Let us assume the actual case of three different behavioral and social groups in Santiago, Chile, A, B, and C, also listed in decreasing order of income and with a high degree of social inequality. Fractions of C are likely to provide services in zones A and B, and therefore interact with their population. Zone B acts as the center of the city, where people meet for office work. The whole dynamics are represented in Fig. 3. Supposing  $e$  is the characteristic fraction of the population working in different zones, the interaction matrix  $\Phi_1$  in normal circumstances (Case 1) is given by Eq. (12). Once the outbreak took place in A, its dynamics were controlled by parameters  $(\beta, \gamma, \theta)$ , which were also valid for its propagation to B. As life quality and access to public health is worse in C, we will suppose such parameters to be, respectively, half and double from the values reported for them in A. The contagious rate in C would also be higher (for example, a 10% higher), as the lack of work-safety policies would press people to work even if they already started to feel sick. Then, the dynamics in C would be characterized by  $(2\beta, \frac{\gamma}{2}, 1.1\theta)$ . Aware of this consideration, but unable to quantify its effect, the government studied to isolate A partially and to limit its contact with C, generating a new interaction matrix  $\Phi_2$  (Case 2):

$$\Phi_1 = \begin{pmatrix} 1 - \frac{e}{2} & \frac{e}{2} & 0 \\ e & 1 - 2e & e \\ e & 2e & 1 - 3e \end{pmatrix}, \quad \Phi_2 = \begin{pmatrix} 1 - \frac{e}{5} & \frac{e}{5} & 0 \\ \frac{e}{5} & 1 - \frac{6e}{5} & e \\ 0 & e & 1 - e \end{pmatrix} \tag{12}$$



**Fig. 3.** Schematic representation of the two interactive configurations between zones A, B, and C. Zone C has a working-dependency with zone A, and zone B acts as the middle point. This is a simplified version of a case observed in Santiago de Chile, where the outbreak was concentrated in the wealthiest part of the city, but quickly spread to zones where service providers live.



**Fig. 4.** Numeric simulations of cases 1-4, with outbreak parameters  $\beta = 0.11$ ,  $\gamma = 0.034$ ,  $\theta = 0.001$ ,  $\xi = 2$ ,  $\alpha = 0.3$ ,  $n = 10^5[2.5\ 5\ 6]$ ,  $\beta^q = 0.044$ ,  $I_A(t=0) = 10$ . Note that this simulation does not consider the saturation of the local health-system, a parameter that drastically increases the mortality rate of this virus when surpassed.

Other studied possibilities were the reduction of social interaction by the application of obligatory quarantine after a time  $t_c$  from the day of the first infection, generating **cases 3 and 4**, which considered the evolution of case 1 or 2 respectively. Following the idea of Del Valle et al. [8], we assume the rate of contagion is directly proportional to the factor of exposure, so that:

$$\frac{\beta^q}{\beta} = \frac{f_{\text{exp}}^1}{f_{\text{exp}}^0}. \quad (13)$$

Fig. 4 show the results of numerical simulations of cases 1-4. The comparison of the different scenarios shows that under these considerations our model was able to adequately represent the effect of the dynamics, providing useful insights to lead the decision-making processes when different behavioral groups interact.

## 5. Conclusions

In the present work, we developed a multi-group SEIRA model that is able to represent the spread of COVID-19 in the SARS-CoV2 outbreak through populations with heterogeneous characteristics, which may be given by geographical particularities of the territory, or marked behavioral differences among social classes in a city, country or region. Because of its generality, this model can represent several mechanisms of interaction between different sub-populations and may lead to a better understanding of the evolution of this pandemic, and thus to better public-health policies.

The general structure of the model presented on this work can be easily modified to account for different particularities of given populations or to support policy makers to take public health decisions with greater effectiveness. For instance, the effect of behavioral changes in the population affecting beta, the saturation of the public health system, and the transient dynamics of the fluxes or interactions between sub-populations can be easily implemented in the model by setting the correct functional form in the parameters. As introducing such dynamics would imply adding more pa-

rameters, we suggest doing so only if there is enough data to fit them.

Since the way individuals interact is one of the principal aspects of human society that has been affected by this worldwide health emergency, we believe that including interactions in our models and analyses will help us develop strategies that will lead us sooner to recover the richness it used to have.

## List of symbols

$X$	Arbitrary variable for representing a generic fraction
$n_i$	Base number of members class $i$
$n_i^e$	Effective number of members class $i$
$\alpha$	Asymptomatic ratio of the population
$\xi$	Extra factor of behavioral virulence of asymptomatic patients
$\Phi^{ij}$	Fraction of class $i$ in class $j$
$p_i$	Immunity ratio of newborns of class $i$
$\Lambda_i$	Net population growth rate $i$
$d_i$	Per-capita base death rate of class $i$
$\beta_i$	Infection rate of the virus in class $i$
$\epsilon_i$	Inverse of the incubation time in class $i$
$\gamma_i$	Recovery rate of class $i$
$\theta_i$	Pathogen induced death rate in class $i$
$\Phi$	Interaction matrix
$f_{\text{exp}}$	Factor of exposure to the infection

## Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## CRediT authorship contribution statement

**Sebastián Contreras:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Supervision. **H. Andrés Villavicencio:** Methodology, Validation, Investigation, Writing - original draft, Writing - review & editing. **David Medina-Ortiz:** Validation, Investigation, Writing - original draft, Writing - review & editing. **Juan Pablo Biron-Lattes:** Validation, Writing - original draft, Writing - review & editing. **Álvaro Olivera-Nappa:** Validation, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

## Acknowledgements

The authors gratefully acknowledge support from the Chilean National Agency for Research and development through ANID PIA Grant AFB180004, and the Centre for Biotechnology and Bioengineering - CeBiB (PIA project FB0001, Conicyt, Chile). DM-O gratefully acknowledges Conicyt, Chile, for PhD fellowship 21181435.

## References

- [1] Aguilar JB, Faust JS, Westafer LM, Gutierrez JB. Investigating the impact of asymptomatic carriers on covid-19 transmission. medRxiv 2020.
- [2] Backer J, Klinkenberg D, Wallinga J. The incubation period of 2019-ncov infections among travellers from wuhan. China medRxiv 2020.
- [3] Bastos SB, Cajueiro DO. Modeling and forecasting the covid-19 pandemic in brazil. arXiv preprint arXiv:200314288 2020.
- [4] Biswas K, Khaleque A, Sen P. Covid-19 spread: Reproduction of data and prediction using a sir model on euclidean network. arXiv preprint arXiv:200307063 2020.
- [5] Calafiore GC, Novara C, Possieri C. A modified sir model for the covid-19 contagion in italy. arXiv preprint arXiv:200314391 2020.
- [6] Chen YC, Lu PE, Chang CS. A time-dependent sir model for covid-19. arXiv preprint arXiv:200300122 2020.
- [7] Day M. Covid-19: identifying and isolating asymptomatic people helped eliminate virus in italian village. BMJ 2020;368:m1165.
- [8] Del Valle SY, Hyman JM, Hethcote HW, Eubank SG. Mixing patterns between age groups in social networks. Social Networks 2007;29:539–54.
- [9] Fanelli D, Piazza F. Analysis and forecast of covid-19 spreading in china, italy and france. Chaos, Solitons & Fractals 2020;134:109761.
- [10] Guo H, Li MY, Shuai Z. Global stability of the endemic equilibrium of multigroup sir epidemic models. Canadian applied mathematics quarterly 2006;14:259–84.
- [11] Hethcote HW, Thieme HR. Stability of the endemic equilibrium in epidemic models with subpopulations. Mathematical Biosciences 1985;75:205–27.
- [12] Kermack WO, McKendrick AG. 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 1927;115:700–21.
- [13] Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM, Sun F, Jit M, Munday JD, et al. Early dynamics of transmission and control of covid-19: a mathematical modelling study. The lancet infectious diseases 2020.
- [14] Lauer SA, Grantz KH. Qifang bi, forrest k jones, qulu zheng, hannah r meredith, andrew s azman, nicholas g reich, and justin lessler. the incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine 2020;3.
- [15] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, et al. Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia. New England Journal of Medicine 2020a.
- [16] Li Y, Shi J, Xia J, Duan J, Chen L, Yu X, Lan W, Ma Q, Wu X, Yuan Y, et al. Clinical features and outcomes of 252 asymptomatic and mildly symptomatic patients with covid-19 in a temporary cabin hospital in wuhan, china. 2020b.
- [17] Mizumoto K, Kagaya K, Zarebski A, Chowell G. Estimating the asymptomatic proportion of coronavirus disease 2019 (covid-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020. Eurosurveillance 2020;25:2000180.
- [18] Pang W. Public health policy: Covid-19 epidemic and seir model with asymptomatic viral carriers. arXiv preprint arXiv:200406311 2020.
- [19] Peng L, Yang W, Zhang D, Zhuge C, Hong L. Epidemic analysis of covid-19 in china by dynamical modeling. arXiv preprint arXiv:200206563 2020.
- [20] Simha A, Prasad RV, Narayana S. A simple stochastic sir model for covid 19 infection dynamics for karnataka: Learning from europe. arXiv preprint arXiv:200311920 2020.
- [21] Sun R, Shi J. Global stability of multigroup epidemic model with group mixing and nonlinear incidence rates. Applied Mathematics and Computation 2011;218:280–6.
- [22] Yang C, Wang J. A mathematical model for the novel coronavirus epidemic in wuhan, china. Mathematical Biosciences and Engineering 2020;17:2708–24.
- [23] Yang Z, Zeng Z, Wang K, Wong SS, Liang W, Zanin M, Liu P, Cao X, Gao Z, Mai Z, et al. Modified seir and AI prediction of the epidemics trend of covid-19 in china under public health interventions. Journal of Thoracic Disease 2020;12:165.
- [24] Zou L, Ruan F, Huang M, Liang L, Huang H, Hong Z, Yu J, Kang M, Song Y, Xia J, et al. Sars-cov-2 viral load in upper respiratory specimens of infected patients. New England Journal of Medicine 2020;382:1177–9.