Epidemiology and Infection

# Supplementary materials for Simulation-based evaluation of school reopening strategies during COVID-19: A case study of São Paulo, Brazil

## E. H. M. CRUZ<sup>1</sup>, J. M. MACIEL<sup>1</sup>, C. CLOZATO<sup>1</sup>, M. S. SERPA<sup>2</sup>, P. O. A. NAVAUX<sup>2</sup>, E. MENESES<sup>3,4</sup>, M. ABDALAH<sup>3</sup>, M. DIENER<sup>5</sup>

<sup>1</sup>Campus Paranavaí, Federal Institute of Parana (IFPR), Paranavaí, Brazil
 <sup>2</sup>Informatics Institute, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil
 <sup>3</sup>Costa Rica National High Technology Center, Costa Rica
 <sup>4</sup>Costa Rica Institute of Technology, Costa Rica
 <sup>5</sup>University of Illinois at Urbana-Champaign, USA

#### This PDF file includes:

Materials and Methods Figs. S1 to S10

Tables S1 to S4

References

#### Other Supplementary Materials for this manuscript include the following:

Repository address of the simulator source code and input data:

https://github.com/ehmcruz/corona

### **1** Materials and Methods

#### **1.1 Generating the Relation Network**

The initial network was created in 5 stages. Note that the initial network considers a normal scenario with no pandemic and interventions. We explain how we handle the interventions in Section 1.2. The initial network was created as follows:

**Home relations** The number of people in the same home follows a distribution described in Table S1 and Figure S1a, which was derived from the microdata of the last official census from the Brazilian Institute of Geography and Statistics (IBGE) [13]. Inside the same home, every person is connected to all other people within the same home.

**Community relations** The number of relations of each person follows a distribution described in Table S1 and Figure S1b.

**Workplace relations** The number of people per company follows a distribution described in Table S1 and Figure S1c. We only model as workplace relations jobs that involve agglomeration of people, such as people that work in the commerce, offices, industry and public sectors. Other jobs, such as people that are autonomous workers or people that work with agriculture, are indirectly modeled in the community relations. The data regarding the number of employed people and per job type was extracted from IBGE [11-13]. In summary, 64% of people with 18 years old or above work. Among people that work, 45% were modeled using workplace relations. Also, inside the same workplace, we consider that the probability of two people to be connected in the network as 50%.

**School relations** The number of students inside the same class and the number of students within each school follow the distributions described in Table S1, Figure S1d and

Figure S1e. For the relations between students from different classrooms, two students have a 0.5% of probability of having a relation in the network. For the relations between students within the same classroom, two students have a 50% of probability of having relation in the network. We add one teacher for each classroom, in which the teacher has a relation to all the students of the same class in the network. According to the census [13], 77.8% of people with 18 years old or below frequent the school.

**Inter-city relations** The number of people that routinely travel between the different cities relations was gathered from the census [13], which indicates a 17.8% of mobility between cities in the metropolitan region of São Paulo. However, the data does not show the cities, only if the person goes to a different city. To overcome this issue, we assume that 90% of the traffic from satellite cities go to the capital, and the rest is divided between the other cities in proportion to their population size. Regarding the traffic from the capital to satellite cities, the traffic is divided proportionally to their population size.

#### **1.2 Handling Interventions**

To handle interventions, we can modify the infection probabilities, as well as add and remove edges from the graph. Different types of intervention require different modifications. We handled the interventions evaluated in the paper as follows:

**Quarantine** To implement the quarantine, we first reduce the infection probabilities inside schools to zero. Afterwards, we remove the workplace relations corresponding to 50% of the workplaces. The 50% of the relations that are left correspond to essential services that can not be interrupted. Finally, we reduce the infection probability of community, workplace and inter-city relations to match the target reduced reproduction rate, which also simulates other sanitary measures such as wearing masks. Home relations are not affected by quarantines. **Reopening of economic activities** During the quarantine, we zeroed the infection probability of half of the workplaces. To reopen the economic activities, we just increase the infection probability of such edges.

**Reopening schools with all students** We only increase the infection probability of school relations, which was zeroed by the quarantine.

**Reopening schools with the São Paulo government strategy** We first remove all school relations from the graph, backing up the data regarding which students and teachers belonged to the same classroom. We create 3 more relation types to represent each group of students. Afterwards, we divide each classroom in 3 and re-create the edges using the newly created relation types. To select which students go to school in a day, we just need to adjust the infection probabilities of the new relations, zeroing the school relations corresponding to the students that will not go to school that day.

**Vaccine** To simulate the vaccines, we only need to change the state of the person being vaccinated to the immune state, carefully considering that it takes some time for the vaccine to make effect, and that not everyone that gets vaccinated will be immunized.

#### **1.3 Other Simulations Parameters**

Several simulation parameters can be found in Table S2 [7, 11-15, 21]. The frequency distributions found in the table can be seen in Figure S2.

In Table S3, we present the probabilities of an infected person to develop symptoms per-age. The biggest challenge to generate this information is that there is no official data regarding the number of unreported cases per age, as expected by the own definition of unreported. Therefore, we need to make several transformations with the reported data to take into account the unreported cases in the simulation. We now describe how we generated these information of Table S3. We first define some functions and variables that return the input data of the algorithm:

- Function *people(age)* returns the number of people of age *age* in the entire population, which we extracted from the census [14].
- Functions reportedMild(age), reportedSevere(age) and reportedCritical(age) should return the number of people with mild, severe and critical symptoms per age, respectively. We extracted this information from the microdata published by the Health Ministry [15].
- Variable *globalRatioUnreported* is the ratio of unreported cases in the entire population. We set it to 0.85 in this paper, since the vast majority of COVID infections are unreported [22]. Although this value seems high, Brazil's underreporting rate is indeed very high [31, 32].

Next, we define some variable and functions to be used later.

$$totalReportedMild = \sum_{age=0}^{9} reportedMild(age)$$
(1)

$$totalReportedSevere = \sum_{age=0}^{9} reportedSevere(age)$$
(2)

$$totalReportedCritical = \sum_{age=0}^{9} reportedCritical(age)$$
(3)

reported(age) = reportedMild(age) + reportedSevere(age) + reportedCritical(age)(4)

$$totalReported = \sum_{age=0}^{9} reported(age)$$
(5)

To calculate the probability of a person to be unreported, we first calculate the ratio of

reported cases:

$$ratioReportedMild(age) = \frac{reportedMild(age)}{reported(age)}$$
(6)

$$ratioReportedSevere(age) = \frac{reportedSevere(age)}{reported(age)}$$
(7)

$$ratioReportedCritical(age) = \frac{reportedCritical(age)}{reported(age)}$$
(8)

Then, we use Equation 10, which finds a value u that forces the mean number of unreported people to the desired value. In the equation, weight(age) returns the proportion of people of age age that should be unreported. Since there is no data for unreported (as expected by the own definition of unreported), we assume that weight has the same ratio of mild cases per age, as it is the softer symptom recorded by official data.

Find *u*, such that:

$$weight(age) = ratioReportedMild(age)$$
(9)

$$u \cdot \frac{\sum_{age=0}^{9} weight(age) \cdot people(age)}{\sum_{age=0}^{9} people(age)} = globalRatioUnreported$$
(10)

After finding u, we calculate a temporary ratio of unreported cases by age, just used to calculate the symptomatic ratios:

$$tmpRatioUnreported(age) = weight(age) \cdot u \tag{11}$$

Now that we know the ratio of unreported cases by age, we can calculate the relative proportion (weight) of symptomatic people:

$$weightMild(age) = ratioReportedMild(age) \cdot (1 - tmpRatioUnreported(age))$$
(12)

$$weightSevere(age) = ratioReportedSevere(age) \cdot (1 - tmpRatioUnreported(age))$$
(13)  
weightCritical(age) = ratioCriticalMild(age) \cdot (1 - tmpRatioUnreported(age)) (14)

Now we define our target averages for the ratio of mild, severe and critical cases relative to the entire population (not only relative to reported cases).

$$targetMildRatio = \frac{totalReportedMild}{totalReported} \cdot (1 - globalRatioUnreported)$$
(15)

$$targetSevereRatio = \frac{totalReportedSevere}{totalReported} \cdot (1 - globalRatioUnreported)$$
(16)

$$targetCriticalRatio = \frac{totalReportedCritical}{totalReported} \cdot (1 - globalRatioUnreported)$$
(17)

Then, we use Equations 18, 19 and 20, which find values m, s and c that force the mean ratio of people to the target value.

Find *m*, such that:

$$m \cdot \frac{\sum_{age=0}^{9} weightMild(age) \cdot people(age)}{\sum_{age=0}^{9} people(age)} = targetMildRatio$$
(18)

Find *s*, such that:

$$s \cdot \frac{\sum_{age=0}^{9} weightSevere(age) \cdot people(age)}{\sum_{age=0}^{9} people(age)} = targetSevereRatio$$
(19)

Find *c*, such that:

$$c \cdot \frac{\sum_{age=0}^{9} weightCritical(age) \cdot people(age)}{\sum_{age=0}^{9} people(age)} = targetCriticalRatio$$
 (20)

Finally, we calculate the ratios relative to the entire population:

$$ratioMild(age) = weightMild(age) \cdot m \tag{21}$$

$$ratioSevere(age) = weightSevere(age) \cdot s \tag{22}$$

$$ratioCritical(age) = weightCritical(age) \cdot c$$
(23)

$$ratioUnreported(age) = 1 - ratioMild(age) - ratioSevere(age) - ratioCritical(age)$$
(24)

Our model allows us to put different infection probabilities depending on the patient symptoms. For patients with severe and critical symptoms, we reduce the infection probabilities by 50% to simulate the more controlled environments inside hospitals. For any symptomatic patients, we zero the infection probabilities to school and workplace relations. For patients with severe and critical symptoms, we zero the infection probabilities to home relations.

Table S4 shows the death probability of each person depending on the symptoms and age.

As we can observe, the age of each person has a deep impact on the symptoms and death rate. Due to that, we use real demographic data from the census [14]. Since the last census is from 2010, we scale the number of people per age such that the total number of people corresponds to the estimated population size in 2019 [14]. The demographic data of the entire São Paulo Metropolitan Area can be seen in Figure S3.

#### **1.4** Calibrating the Simulation to Match Real-World Behavior

Figure S4 contains a comparison of the simulator behavior to the real-world data of the São Paulo Metropolitan Area. The calibration process involves setting the interventions, as explained in Section 1.2, as well as manipulating the infection probabilities. Calibration was performed for the first 210 days of the pandemic, starting at 26 February 2020, which is the date corresponding to the first documented case in Brazil. In Figure S4a, we compare the estimated number of reported people in the infected state at a day (do not confuse with the number of newly infected people per day). In Figure S4b, we compare the accumulated number of reported people in the infected state. For these two calibrations, we used the value from Equation 5, thereby unreported people were not considered, since they are not part of official statistics. In Figure S4c, we compare the number of critical patients, which represents the ICU occupancy. We also show a comparison of per-age statistics in Figure S5, where we can observe that the simulator results follow the same tendency of real data, with a concentration of reported cases in adults and a concentration of deaths in elder people. It is important to mention that a perfect match between simulator and real-world behavior is impossible, mostly because of the fact that the vast majority of cases are unreported [22], such that it is impossible to accurately measure real behavior with the official published data. Nevertheless, we were able to configure the simulation parameters to reflect the real behavior with a reasonable precision, as can be seen in the figures.

Figs. S1 to S10



Fig. S1. Frequency distributions used to create the network of relations of the population.



Fig. S2. Frequency distributions of the parameters of Table S2.



Fig. S3. Demographic data of the population of the São Paulo Metropolitan Area.



Fig. S4. Comparison of simulator behavior to the real-world behavior.



(e) Amount of deaths per age (simulator).

Fig. S5. Per-age amount of COVID-19 cases and deaths, from real data and from the simulator (using the São Paulo plan experiment).



Fig. S6. Curves of infected people when schools reopen with all students at once.



Fig. S7. Curves of infected people when schools reopen following São Paulo's strategy.



Fig. S8. Curves of infected people when schools reopen following São Paulo's strategy, evaluating different isolation levels.



(a) All students at once – accumulated reported cases.



Fig. S9. Curves of infected people considering a lower asymptomatic transmissibility.

Schools closed Reopen – best case Reopen – worst case

schools reopen

90 180 270 360 450 540 630 720

Days since start of pandemic

(b) São Paulo's strategy - accumulated re-

1,560,000

1,300,000

1,040,000

780,000

520,000

260,000

ported cases.

0

Ò

Number of cases



cases.



Fig. S10. Curves of infected people when schools reopen with all students on day 400.

Schools closed Reopen – best case Reopen – worst case

vaccination starts

720,000

600,000

480,000 360,000

240.000

120,000

0

Ò

Number of cases

schools reopen

~

90 180 270 360 450 540 630 720

# Tables S1 to S4

Relation	Description	Туре	Mean	StdDev	Min	Max
Home	Number of people per home	Gamma	3.3	1.7	1	10
Community	Number of community relations per person	Gamma	15.0	10.0	5	50
Workplace	Number of people per company	Gamma	8.0	7.0	2	50
School – classroom	Number of people per classroom	Gamma	30.0	10.0	15	50
School – total	Number of people per school	Gamma	500.0	500.0	300	4,000

 Table S1. Frequency distributions used to generate the relation network.

Parameter	Value		
Cycle length	1 day.		
$R_0$ (basic reproduction rate)	3.0 (at the start of the pandemic).		
Pre-symptomatic time	1 day.		
Time for which people with unreported			
and mild symptoms are contagious	4 days.		
Incubation time	Gamma, mean: 4.6 days, Stddev: 3.2, Min: 1, Max: 14.		
Time from symptomatic to hospitalization	5 days.		
Days in ICU for critical patients	Gamma, mean: 15.0 days, Stddev: 15.0, Min: 1, Max: 60.		
Days in hospital for severe patients (infirmary)	Gamma, mean: 6.5 days, Stddev: 5.5, Min: 1, Max: 60.		
Age of the teachers	Gamma, mean: 41.2 years, Stddev: 9.9, Min: 20, Max: 70.		
Vaccine – time to be effective	14 days.		
Vaccine – number of people vaccinated per day	300,000 people per day.		
Vaccine – immunity probability	80%.		
Infection probability – home	$\times 3.0$ the community probability considering the begin ning of the pandemic (not changed by any intervention)		
Infection probability – workplace	$\times 1.5$ the community probability.		
Infection probability – school	We evaluated both $\times 2$ and $\times 4$ the community probability.		
Infection probability – inter-city	Equal to the community probability.		
Infection probability – community	We set it such that the epidemic behavior matches the target reproduction rate.		

 Table S2. Other simulation parameters.

Age	Unreported	Mild	Severe	Critical
0-10	0.8916	0.1051	0.0024	0.0008
10-20	0.9613	0.0384	0.0003	0.0001
20-30	0.9447	0.0545	0.0006	0.0002
30-40	0.8987	0.0985	0.0022	0.0006
40-50	0.8340	0.1577	0.0064	0.0019
50-60	0.7211	0.2513	0.0202	0.0074
60-70	0.5560	0.3379	0.0737	0.0323
70-80	0.5079	0.1459	0.2274	0.1188
80-90	0.4939	0.2276	0.1854	0.0930
90+	0.4931	0.2283	0.1901	0.0886
Mean	0.8500	0.1235	0.0182	0.0084

Table S3. Probability of an infected person to develop specific symptoms depending on the age.

Age	Death (severe cases)	Death (critical cases)
0-10	0.0093	0.0649
10-20	0.0264	0.1405
20-30	0.0194	0.1545
30-40	0.0312	0.1601
40-50	0.0517	0.2214
50-60	0.0932	0.3086
60-70	0.1623	0.4349
70-80	0.2379	0.5128
80-90	0.3548	0.5549
90+	0.4702	0.5576
Mean	0.1365	0.3742

Table S4. Probability of a person to die depending on their symptoms and age.

### References

- 1. C. C. Kerr, et al., Covasim: an agent-based model of COVID-19 dynamics and interventions, medRxiv (2020).
- M. Worobey, et al., The emergence of SARS-CoV-2 in Europe and North America, Science eabc8169 (2020).
- L. G. Barberia, E. J. Gómez, Political and institutional perils of Brazil's COVID-19 crisis, The Lancet 396, 367 (2020).
- D. S. Hui, et al., The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health — The latest 2019 novel coronavirus outbreak in Wuhan, China, International Journal of Infectious Diseases 91, 264 (2020).
- 5. W. Kawohl, C. Nordt, *COVID-19, unemployment, and suicide, The Lancet Psychiatry* 7, 389 (2020).
- 6. I. F. Miller, A. D. Becker, B. T. Grenfell, C. J. E. Metcalf, *Disease and healthcare burden* of COVID-19 in the United States, Nature Medicine **26**, 1212 1217 (2020).
- 7. J. Phua, et al., Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations, The Lancet Respiratory Medicine **8**, 506 (2020).
- N. M. Ferguson, *et al.*, Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand, *Tech. rep.*, Imperial College London (2020).
- T. Mellan, *et al.*, Report 21: Estimating covid-19 cases and reproduction number in brazil, *Tech. rep.*, Imperial College London (2020).
- 10. J. Enright, R. R. Kao, Epidemics on dynamic networks, Epidemics 24, 88 (2018).

- IBGE, Indicadores ibge pesquisa nacional por amostra de domicílios contínua. trimestre móvel out. - dez. 2019, *Tech. rep.*, Instituto Brasileiro de Geografia e Estatistica - Brazil (2020).
- 12. IBGE, Pnad contínua: taxa de desocupação é de 11.0% e taxa de subutilização é de 23.0% no trimestre encerrado em dezembro. https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/26740-pnad-continua-taxa-de-desocupacao-e-de-11-0-e-taxa-de-subutilizacao-e-de-23-0-no-trimestre-encerrado-em-dezembro, instituto brasileiro de geografia e estatistica brazil, [accessed sept. 16th, 2020] (2020).
- IBGE, Censo 2010 microdados. https://www.ibge.gov.br/estatisticas/sociais/populacao/ 22827-censo-2020-censo4.html?=&t=microdados, instituto brasileiro de geografia e estatistica - brazil, [accessed sept. 16th, 2020] (2010).
- SUS, Datasus portal da saúde. http://www2.datasus.gov.br/DATASUS/index.php, sistema Único de saúde, ministério da saúde, [accessed sept. 16th, 2020] (2020).
- SUS, Sus opendata. https://opendatasus.saude.gov.br/, sistema Único de saúde, ministério da saúde, [accessed sept. 16th, 2020] (2020).
- 16. SEADE, Repositório de dados sobre casos e óbitos decorrentes do covid-19 nos municípios do estado de são paulo e sobre leitos e internações por departamento regional de saúde; https://github.com/seade-R/dados-covid-sp, seade, secretaria de estado da saúde de são paulo, [accessed sept. 16th, 2020] (2020).
- 17. T. Britton, F. Ball, P. Trapman, A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2, Science **369**, 846 (2020).

- R. Lordan, G. A. FitzGerald, T. Grosser, *Reopening schools during COVID-19*, *Science* 369, 1146 (2020).
- 19. C. J. E. Metcalf, D. H. Morris, S. W. Park, *Mathematical models to guide pandemic response*, *Science* **369**, 368 (2020).
- 20. A. Abou-Ismail, Compartmental Models of the COVID-19 Pandemic for Physicians and Physician-Scientists, SN Comprehensive Clinical Medicine 2, 852 (2020).
- 21. E. M. Rees, et al., COVID-19 length of hospital stay: a systematic review and data synthesis, BMC Medicine 18, 270 (2020).
- 22. R. Li, et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2), Science **368**, 489 (2020).
- 23. J. Panovska-Griffiths, et al., Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study, The Lancet Child and Adolescent Health (2020).
- SP, Plano de Retorno da Educação Plano São Paulo Governo do Estado de São Paulo. https://www.saopaulo.sp.gov.br/planosp/, [accessed Sept. 21st, 2020] (2020).
- F. Brauer, Mathematical epidemiology: Past, present, and future, Infectious Disease Modelling 2 (2017).
- 26. M. Graham, T. House, *Dynamics of stochastic epidemics on heterogeneous networks*, *Journal of mathematical biology* **68** (2013).
- 27. T. L. Psychiatry, Isolation and inclusion, The Lancet Psychiatry 7, 371 (2020).

- R. M. Viner, et al., School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review, The Lancet Child and Adolescent Health 4, 397 (2020).
- 29. H. Hildenwall, et al., Paediatric COVID-19 admissions in a region with open schools during the two first months of the pandemic, Acta Paediatrica **109**, 2152 (2020).
- S. H. Vermund, V. E. Pitzer, Asymptomatic transmission and the infection fatality risk for COVID-19: Implications for school reopening, Clinical Infectious Diseases (2020). Ciaa855.
- W. P. Sobrinho, Casos de covid-19 são 10 vezes o notificado e somam 7 milhões, diz estudo – https://noticias.uol.com.br/saude/ultimas-noticias/redacao/2020/06/09/casosde-covid-19-sao-10-vezes-o-notificado-com-6-mi-de-doentes-diz-estudo.htm, UOL [accessed Dec. 21st, 2020] (2020).
- 32. C. Boehm, Pesquisa da UFPel estima subnotificação de casos de covid-19 no Brasil – https://agenciabrasil.ebc.com.br/educacao/noticia/2020-06/pesquisa-da-ufpelestima-subnotificacao-de-casos-de-covid-19-no-brasil, Agencia Brasil [accessed Dec. 21st, 2020] (2020).
- 33. Globo, SP aplica mais de 358 mil doses de vacina contra a gripe no primeiro dia da campanha – https://g1.globo.com/sp/sao-paulo/noticia/2020/03/24/sp-aplica-maisde-306-mil-doses-de-vacina-contra-a-gripe-no-primeiro-dia-da-campanha.ghtml, G1 [accessed Dec. 21st, 2020] (2020).
- 34. I. Dattner, et al., The role of children in the spread of COVID-19: Using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children, medRxiv (2020).

- 35. D. He, et al., The relative transmissibility of asymptomatic COVID-19 infections among close contacts, International Journal of Infectious Diseases **94**, 145 (2020).
- P. H. A. of Sweden, Covid-19 in schoolchildren a comparison between finland and sweden, *Tech. rep.*, Public Health Agency of Sweden (2020).
- 37. C. M. D. MELO, G. A. SILVA, A. R. MELO, A. C. D. FREITAS, *COVID-19 pandemic outbreak: the Brazilian reality from the first case to the collapse of health services, Anais da Academia Brasileira de Ciencias* **92** (2020).