# Travel Behavior and Individual Choices in Infectious Disease Spread: Enhancing Activity-Based Models with Awareness and Testing Dynamics

Cloe Cortes Balcells[1] and Michel Bierlaire[1]

August 15, 2024

**Keywords:** Activity-Travel Behavior, Epidemiological Models, Individual Behaviors, Decision Making, Public Health Strategies, Large-Scale Simulations.

## 1 Introduction

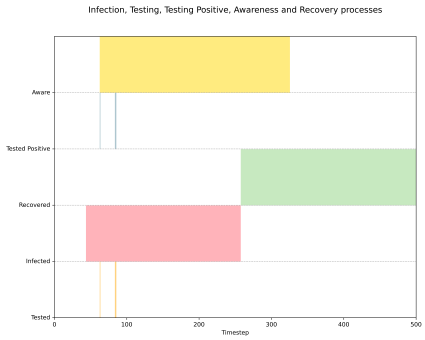
Beyond impacting global economies and healthcare structures, the COVID-19 pandemic has revealed several weaknesses in the ability of authorities to handle large-scale crises, notably in understanding the evolution of people's activity-travel behavior during such crises [1]. How individuals move, interact, and make health-related decisions (e.g. undergoing testing or voluntarily limiting their trips) are key to understanding the propagation of infectious diseases. Therefore, accounting for these dynamics is necessary to implement efficient and targeted policies to contain disease outbreaks. In this context, epidemiological models serve as invaluable tools for formulating effective interventions.

Aggregated and disaggregated epidemiological models exist and can be classified into three groups: compartmental [2], network-based [3, 4], and activity-based [5, 6]. Compartmental models group the population into different categories, neglecting the heterogeneity of the individuals that are part of it. Network-based models, on the other hand, delve into the interactions between individuals within interconnected networks, but are too computationally expensive to apply for large-scale populations. In this study, we focus on activity-based models, which present a detailed representation of individual behaviors. These models are especially valuable for understanding disease spread in complex real-world scenarios [5, 7]. However, while activity-based epidemiological models have been widely used to guide public health responses during the pandemic [5], they have some limitations.

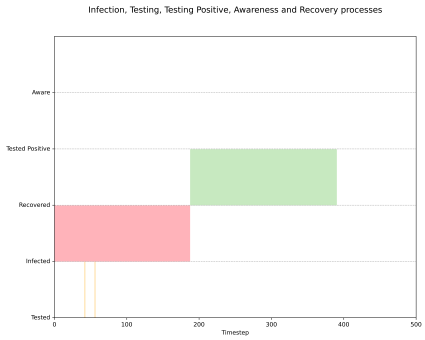
Existing models [5, 8, 9, 10], typically assign a uniform probability of infection to individuals or make adjustments primarily based on age, overlooking the influence of health characteristics and the attributes of infectious contacts encountered. Furthermore, the literature primarily focuses on the implications of activity restrictions [11], leaving a gap in the exploration of the effects of various policies. For example, there is a notable lack of studies that assess the impact of policies such as requiring negative tests to participate in leisure activities. Lastly, models generally emphasize the infection process [6, 5], often neglecting individual choices, such as someone who decides to get tested. This approach provides a limited view of disease propagation and individual responses to infection.

Our model aims to address these gaps by creating a detailed epidemiological model that captures both the dynamics of infection and the results of the tests on an individual level. Both

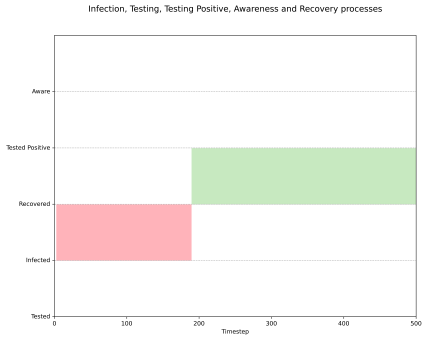
[1]Ecole Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, cloe.cortesbalcells@epfl.ch



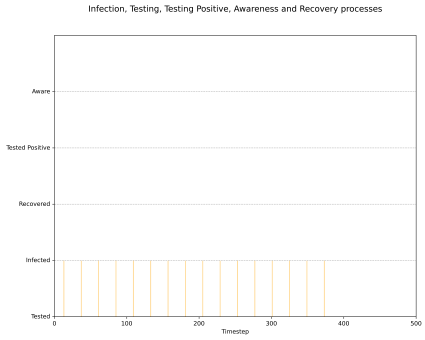

(a) Aware and Responsible




(b) False Hope




(c) Blissful Ignorance




(d) Anxiety and Vigilance

Figure 1: Representation of individuals' awareness and response patterns towards infection and testing behaviors. (a) Aware and Responsible: An individual aware of her infection, adjusting behavior to minimize transmission risk. (b) False hope: An individual with a false negative test who, unknowingly, carries and potentially spreads the infection. (c) Blissful Ignorance: An individual unaware and untested, highlighting risks of lack of testing or misinformation. (d) Anxiety and Vigilance: An individual frequently tests due to high anxiety, usually receiving negative results.

models account for latent agent behavior. By simulating behaviors over discrete time intervals, the model facilitates in-depth exploration of transitions in health states, changes in infection awareness, and the effects of testing. The central components of the model are: (i) movement of individuals throughout the different facilities, (ii) changes in individuals' health state, (iii) individuals choice to get tested, and (iv) individual changes in activity-travel behavior given the test outcomes. The concept of "awareness" is an important factor in an individual's disease journey. Awareness is a binary variable that captures the moment when individuals become aware of their infection by testing positive. This information is critical, as awareness directly influences mobility and activity-travel behavior: individuals change their daily schedules and interactions only when they realize their infection status. An example is provided in Figure 1 illustrating the behaviors of four different agents within our model. By incorporating these sub-models, our model offers a comprehensive perspective on the relationship between the choices of individuals, the spread of the disease, and their activity-travel behavior. Moreover, we can test a range of policies beyond activity restrictions, and assess how "responsible" individuals are, i.e. if they are willing to compile with testing policies. Finally, our model presents very high computational efficiency, making it a tool for large-scale simulations and real-time public health decision-making.

# 2 Methodology

## 2.1 Data

The dynamics of disease spread occur in $t$ periods. For mobility and agent characteristics, we use the output of a microscopic activity-based model [12]. The model describes a population consisting of $N$ individuals who visit the facilities $f$ from the set $\mathcal{F}$, and have socioeconomic characteristics $x_n^o$, and health characteristics $x_n^h$. For each individual $n$, facility $f$, and timestep $t$, we denote $Z_{fnt}^e$ as a binary variable that:

$$Z_{fnt}^e = \begin{cases} 1 & \text{if individual } n \text{ is in } f \text{ at time } t \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Regarding infection data, use data from the Federal Office of Public Health (FOPH) on daily positive tests and information on individuals tested in Switzerland from mid-February 2020 to mid-September 2021 [13]. This dataset includes age, sex, municipality, vaccination doses, hospitalization, and causalities for all the individuals who tested positive. Additionally, we use open-source aggregated data [14] from Switzerland, including positive, negative, and tested counts, per age group. This dataset is required since we need information on the total tests. The segmentation of the population is done into 10 age groups $g$. The group $g$ to which an individual $n$ belongs is denoted as $g_n$. The initial conditions of the framework have an impact on the simulation results, as the trajectory of an epidemic is directly related to the number of initial cases. We assume the model to run after ten days from the start of the pandemic. For this reason, the initial conditions for the first days of simulation are not determined by the model, but are taken directly from the data [14].

## 2.2 Modeling Elements

- **Contacts:** We define a contact as an interaction between individuals where there is the potential for the transmission of the disease. The spread of the disease is modeled on the basis of the contacts generated by the activity schedules of the population from the microscopic activity-based model. We define $Z_{nmt}^c$ as a binary variable indicating if individual $n$ and individual $m$ meet at timestep $t$.

$$Z_{nmt}^c = \sum_f Z_{fnt}^e \cdot Z_{fmt}^e, \quad \forall m \in \mathcal{N} \backslash \{n\},$$

  where the product between $Z_{fnt}^e$ and $Z_{fmt}^e$ is 1 only if both $n$ and $m$ are in facility $f$ at the same timestep $t$.

- **Health State:** The health state of every individual during any period is classified as susceptible, infected, or recovered. We use three binary variables to define the health state: $Z_{nt}^s$, $Z_{nt}^i$ and $Z_{nt}^r$; where $Z_{nt}^s = 1$ if the health state of the individual is "susceptible", $Z_{nt}^i = 1$ if the individual is "infected", and $Z_{nt}^r = 1$ if the individual is "recovered", such that:

$$Z_{nt}^s + Z_{nt}^i + Z_{nt}^r = 1 \quad \forall n, t. \tag{2}$$

- **Model Outcomes:** The model emphasizes the movement of individuals between health states and their responses to test outcomes. A latent state called "exposure" indicates

an individual's infection likelihood. The indicators that we obtain for each individual at every timestep $t$ are defined in Table 1. It includes mobility indicators: $Z^e_{nt}$ and $Z^c_{nmt}$, latent state of exposure $E^\star_{nt}$, behavioral indicators: $Z^a_{nt}$ and $Z^q_{nt}$, and epidemiological: $Z^s_{nt}$,$Z^i_{nt}$,$Z^r_{nt}$, and $Z^+_{nt}$.

- **Recovery Time:** This represents the time it takes for an individual to transition from an infected state to a recovered state. It is assumed to follow a log-normal distribution as in [15].

$$\gamma_n \sim \text{lognormal}(384, 96), \quad \forall n, \tag{3}$$

where the mean and standard deviation are expressed in time steps $t$.

| Name | Description | Type |
|---|---|---|
| $E^\star_{nt}$ | latent state of exposure | continuous |
| $Z^s_{nt}$ | the individual is susceptible | binary |
| $Z^i_{nt}$ | the individual is infected | binary |
| $Z^r_{nt}$ | the individual is recovered | binary |
| $Z^+_{nt}$ | testing results of each individual | binary |
| $Z^a_{nt}$ | awareness of each individual | binary |
| $Z^q_{nt}$ | propensity of test | binary |
| $Z^e_{nt}$ | location of the individual | binary |
| $Z^c_{nmt}$ | contact between two individuals | binary |

Table 1: Indicators epidemiological model for every individual $n$ throughout $p$.

## 2.3 Model

Operating within discrete time intervals, the model captures the intricacies of infection spread, testing procedures, and the resulting behavioral changes. A graphical representation of the dynamics of the model can be found in Figure 2. We summarize the dynamics of the model in the following steps. First, we update time and agent locations based on daily activities (ABM) (see Equation 1).

After updating the time and agent locations, we update agent health states following the state transition matrix $\mathbb{B}$ that contains the probabilities of transitioning from one state: susceptible, infected and recovered, to another state: susceptible, infected, and recovered, at any time $t$.

$$\mathbb{B} = \begin{bmatrix} P(Z^s_{n(t+1)} = 1 | Z^s_{nt} = 1) & P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) & 0 \\ 0 & P(Z^i_{n(t+1)} = 1 | Z^i_{nt} = 1) & P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) \\ P(Z^s_{n(t+1)} = 1 | Z^r_{nt} = 1) & 0 & P(Z^r_{n(t+1)} = 1 | Z^r_{nt} = 1) \end{bmatrix}$$

where:

$$P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1) = \frac{1}{1 + e^{-\mu E^\star_{nt}}}, \tag{4}$$

$$P(Z^s_{n(t+1)} = 1 | Z^s_{nt} = 1) = 1 - P(Z^i_{n(t+1)} = 1 | Z^s_{nt} = 1), \tag{5}$$

$$P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) = \Phi_{\gamma_n}(p - p^i_n) \tag{6}$$

$$P(Z^i_{n(t+1)} = 1 | Z^i_{nt} = 1) = 1 - P(Z^r_{n(t+1)} = 1 | Z^i_{nt} = 1) \tag{7}$$

$$P(Z^r_{n(t+1)} = 1 | Z^r_{nt} = 1) = 0 \tag{8}$$

$$P(Z^s_{n(t+1)} = 1 | Z^r_{nt} = 1) = 1 \tag{9}$$

The hidden exposure state $E^\star_{nt}$ is a continuous variable that captures the individual's level of exposure to infection. This latent state takes as exogenous variables the health characteristics of the individual $x^h_n$, and the number of infectious contacts $\chi^i_{nt}$ for an individual $n$ for timestep $t$.

$$E^*_{nt} = \beta^e_0 + \sum_{k=1}^{K_e - 1} \beta^h_k x^h_{nk} + \beta^i \chi^i_{fnt} + \varepsilon^e, \tag{10}$$

where $\beta^h_k$ is a vector of $K_e$ parameters (to be estimated from the data), $\beta^i$ is the parameter of $\chi^i_{fnt}$, $\varepsilon^e$ is the (random) error term, and $\chi^i_{fnt}$ is defined as:

$$\chi^i_{fnt} = \frac{\sum_m Z^i_{mt} Z^e_{fmt}}{\sum_m Z^e_{fmt}}. \tag{11}$$

Then, we run the testing model to estimate the binary choice of getting a test or not for every individual. We estimate the probability for an individual to perform a test in timestep $t$ as:

$$P(Z^q_{nt} = 1) = \frac{1}{1 + e^{-\mu(\beta^q_0 + \sum_{k=1}^{K_q - 1} \beta^o_k x^o_{nk} + \eta^e E^\star_{nt})}} \tag{12}$$

where $\beta^o_k$ is the parameter vector of $K_q$ parameters (to be estimated from data) from the socio-economic characteristics of the individual $x^o_n$, and $\eta^e$ is the parameter for $E^\star_{nt}$. Once we know which individuals get tested, we estimate the number of positive individuals, by first checking whether the individual is selected for testing, i.e. if

$$Z^q_{nt} = 1, \tag{13}$$

defined in Table 1. After, we divide all individuals who fulfill (13), according to their health status as:

$$n^{qi}_t \in Z^q_{nt} = 1 \ \& \ Z^i_{nt} = 1, \tag{14}$$

$$n^{qs}_t \in Z^q_{nt} = 1 \ \& \ Z^s_{nt} = 1, \tag{15}$$

$$n^{qr}_t \in Z^q_{nt} = 1 \ \& \ Z^r_{nt} = 1. \tag{16}$$

where $n^{qi}_t$ is an individual that is infected and tested, at time $t$, $n^{qs}_t$ is an individual that is susceptible and tested, and $n^{qr}_t$ is an individual that is recovered and tested. Additionally, we define the probabilities of PCR sensitivity as a constant value throughout the simulation, in $n$ and $p$ [16]. The ranges in parentheses correspond to the confidence intervals $95\%$.

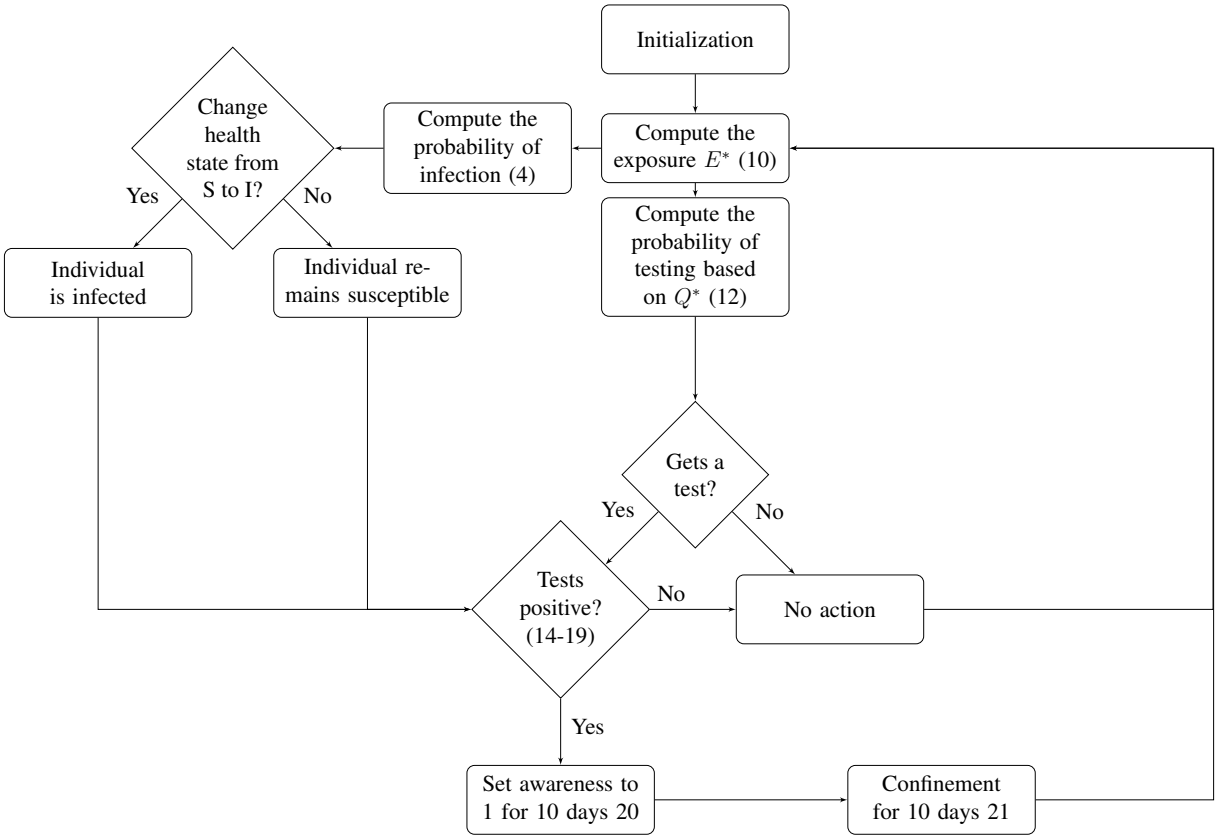



Figure 2: Overall dynamics of the discrete choice models inside the framework

$$P(Z^+_{nt} = 1 | Z^q_{nt} = 1 \text{ and } Z^i_{nt} = 1) = 0.65 \pm (0.62 - 0.68), \tag{17}$$
$$P(Z^+_{nt} = 1 | Z^q_{nt} = 1 \text{ and } Z^s_{nt} = 1) = 0.17 \pm (0.10 - 0.23), \tag{18}$$
$$P(Z^+_{nt} = 1 | Z^q_{nt} = 1 \text{ and } Z^r_{nt} = 1) = 0.17 \pm (0.10 - 0.23). \tag{19}$$

Since our hypothesis is that individuals only change their behavior once they are aware of being sick, that is $Z^i_{nt} = 1$ and $Z^q_{nt} = 1$. Therefore, we compute the individual awareness of infection indicator which is modeled as:

$$Z^a_{nt} = Z^i_{nt} Z^q_{nt}. \tag{20}$$

Finally, we implement the mobility restrictions by:

$$Z^e_{fn(t+1)} = \begin{cases} Z^e_{fnt} & \text{if individual's } n \text{ outcome is 0, and} \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

# 3 Results

We calibrate our model using the data described in Section 2.1, applying measurement equations that link individual tests with observed tests per population segment each week. To deal with the aggregation of the data and keep the individual-level information of the model, the calibration of the model parameters uses the negative binomial distribution, chosen for its precision in

depicting variability in test data. However, adjusting it directly is difficult due to its distinct and non-smooth nature. The details of this methodology are beyond the scope of this extended abstract but will be fully discussed in future work.

One of the standout features of our model is its computational efficiency; we can process 800,000 individuals in about 3 seconds over a three-month simulation, making it a very suitable tool for large-scale scenarios.

Figure 1 represents the different responses to infection and testing behaviors. Each subfigure highlights a distinct type of behavioral attitude of an individual based on their awareness and the changes in their activity-travel behavior. Figure 1a shows an individual who is aware of her infection and actively modifies her behavior during travel activity to minimize transmission risk, representing the ideal behavioral response where knowledge leads to responsible actions, with indicators: $Z^i_{nt} = 1, Z^q_{nt} = 1, Z^+_{nt} = 1$, and $Z^a_{nt} = 1$. Following this, in Figure 1b we observe an individual testing with a false negative, and therefore becoming an unaware carrier of the infection, potentially endangering others. In contrast, Figure 1c portrays an individual who is oblivious to its infection. She chooses not to get a test and therefore can inadvertently spread the disease, highlighting the risks associated with a lack of testing or misinformation. Lastly, Figure 1d, presents an individual characterized by high anxiety about contracting the disease, frequently testing, and getting negative tests.

To examine our model's accuracy and applicability, we tested it on several data sets featuring various demographic and mobility patterns. Preliminary results indicate substantial variations in disease spread due to differences in individual awareness and testing tendencies. Specifically, areas with higher awareness and testing experienced more controlled spread, underscoring the importance of public health measures and individual actions.

# 4 Conclusion

In this extended abstract, we uncover a critical oversight in traditional disease modeling. Many models are designed under the assumption that a positive test result directly equates to an individual having an active infection, which does not account for the different behavioral responses of individuals who are unaware of their infection status. Our research emphasizes the importance of distinguishing between actual infection and an individual's awareness of it. This differentiation is crucial. It helps us differentiate people who remain oblivious to their infection and thus do not change their behavior, to those who become cautious upon testing positive and change their activity-travel behavior. Moreover, we can capture other behaviors like individuals who frequently test out of fear without ever getting a positive result. By understanding these nuances, we get a clearer picture of how diseases spread in the context of an individual's choices and activity-travel behavior. This insight can be particularly helpful for predicting how diseases impact daily activities such as travel.

Also, the computational efficiency not only makes it a valuable tool for academic research, but also a practical resource for real-time decision-making by those in charge of public health responses.

Finally, our approach draws from various disciplines, such as epidemiology, transportation, and discrete choice analyses. By merging these fields, we have developed a comprehensive approach to understanding the spread of disease. This model not only sheds light on the SARS-CoV-2 pandemic but also sets the stage for how we can better approach future infectious diseases.

# References

[1] J. T. Tuomisto, J. Yrjola, M. Kolehmainen, J. Bonsdorff, J. Pekkanen, and T. Tikkanen, "An agent-based epidemic model REINA for COVID-19 to identify destructive policies," Infectious Diseases (except HIV/AIDS), preprint, Apr. 2020. [Online]. Available: http://medrxiv.org/lookup/doi/10.1101/2020.04.09.20047498

[2] W. O. Kermack, A. G. McKendrick, and G. T. Walker, "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*, vol. 115, no. 772, pp. 700–721, Aug. 1927, publisher: Royal Society. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118

[3] M. Mancastroppa, R. Burioni, V. Colizza, and A. Vezzani, "Active and inactive quarantine in epidemic spreading on adaptive activity-driven networks," *Physical Review E*, vol. 102, no. 2, p. 020301, Aug. 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.102.020301

[4] S. Eubank, H. Guclu, V. S. Anil Kumar, M. V. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang, "Modelling disease outbreaks in realistic urban social networks," *Nature*, vol. 429, no. 6988, pp. 180–184, May 2004, number: 6988 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/nature02541

[5] C. C. Kerr, R. M. Stuart, D. Mistry, R. G. Abeysuriya, G. Hart, K. Rosenfeld, P. Selvaraj, R. C. Nunez, B. Hagedorn, L. George, A. Izzo, A. Palmer, D. Delport, C. Bennette, B. Wagner, S. Chang, J. A. Cohen, J. Panovska Griffiths, M. Jastrzebski, A. P. Oron, E. Wenger, M. Famulare, and D. J. Klein, "Covasim: an agent-based model of COVID-19 dynamics and interventions," May 2020, pages: 2020.05.10.20097469. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.05.10.20097469v1

[6] S. A. Muller, M. Balmer, B. Charlton, R. Ewert, A. Neumann, C. Rakow, T. Schlenther, and K. Nagel, "Using mobile phone data for epidemiological simulations of lockdowns: government interventions, behavioral changes, and resulting changes of reinfections," Epidemiology, preprint, Jul. 2020. [Online]. Available: http://medrxiv.org/lookup/doi/10.1101/2020.07.22.20160093

[7] A. Tirachini and O. Cats, "COVID-19 and Public Transportation: Current Assessment, Prospects, and Research Needs," *Journal of Public Transportation*, vol. 22, no. 1, Jan. 2020. [Online]. Available: https://scholarcommons.usf.edu/jpt/vol22/iss1/1

[8] K. Nagel, C. Rakow, and S. A. Müller, "Realistic agent-based simulation of infection dynamics and percolation," *Physica A: Statistical Mechanics and its Applications*, vol. 584, p. 126322, Dec. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378437121005951

[9] M. E. Halloran, N. M. Ferguson, S. Eubank, I. M. Longini, D. A. T. Cummings, B. Lewis, S. Xu, C. Fraser, A. Vullikanti, T. C. Germann, D. Wagener, R. Beckman, K. Kadau, C. Barrett, C. A. Macken, D. S. Burke, and P. Cooley, "Modeling targeted layered containment of an influenza pandemic in the United States," *Proceedings of the National Academy of Sciences*, vol. 105, no. 12, pp. 4639–4644, Mar. 2008. [Online]. Available: https://pnas.org/doi/full/10.1073/pnas.0706849105

[10] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani, "Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand," Imperial College London, Tech. Rep., Mar. 2020. [Online]. Available: http://spiral.imperial.ac.uk/handle/10044/1/77482

[11] A. Aleta, D. Martín-Corral, M. A. Bakker, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini, A. Pentland, A. Vespignani, Y. Moreno, and E. Moro, "Quantifying the importance and location of SARS-CoV-2 transmission events in large metropolitan areas," Epidemiology, preprint, Dec. 2020. [Online]. Available: http://medrxiv.org/lookup/doi/10.1101/2020.12.15.20248273

[12] S. Horl and M. Balac, "Synthetic population and travel demand for Paris and Ile-de-France based on open and publicly available data," *Transportation Research Part C: Emerging Technologies*, vol. 130, p. 103291, Sep. 2021. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0968090X21003016

[13] J. Riou, R. Panczak, C. L. Althaus, C. Junker, D. Perisa, K. Schneider, N. G. Criscuolo, N. Low, and M. Egger, "Socioeconomic position and the COVID-19 care cascade from testing to mortality in Switzerland: a population-based analysis," *The Lancet Public Health*, vol. 6, no. 9, pp. e683–e691, Sep. 2021, publisher: Elsevier. [Online]. Available: https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(21)00160-2/fulltext

[14] G. CloudPlatform, "Google Covid data," 2021. [Online]. Available: https://github.com/GoogleCloudPlatform/covid-19-open-data/blob/main/docs/table-epidemiology.md

[15] R. Wolfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brunink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, and C. Wendtner, "Virological assessment of hospitalized patients with COVID-2019," *Nature*, vol. 581, no. 7809, pp. 465–469, May 2020, number: 7809 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41586-020-2196-x

[16] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, "Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, Aug. 2020. [Online]. Available: http://pubs.rsna.org/doi/10.1148/radiol.2020200642