Sustainable targeted interventions to mitigate the COVID-19 pandemic: A big data-driven modeling study in Hong Kong

Hanchu Zhou, Qingpeng Zhang *Senior Member, IEEE*, Zhidong Cao, Helai Huang, and Daniel Dajun Zeng, *Fellow, IEEE*

***Abstract*— Nonpharmaceutical interventions (NPIs) for contact suppression have been widely used worldwide, which impose harmful burdens on the population and the local economy. The evaluation of alternative NPIs is needed to confront the pandemic with less disruption. By harnessing human mobility data, we develop an agent-based model that can evaluate the efficacies of NPIs with individualized mobility simulations. Based on the model, we propose data-driven targeted interventions to mitigate the COVID-19 pandemic in Hong Kong without city-wide NPIs. We develop a data-driven agent-based model for 7.55 million Hong Kong residents to evaluate the efficacies of various NPIs in the first 80 days of the initial outbreak. The entire territory of Hong Kong is split into 4,905 500m $\times$ 500m grids. The model can simulate detailed agent interactions based on the demographics data, public facilities and functional buildings, transportation systems, and travel patterns. The general daily human mobility patterns are adopted from Google's Community Mobility Report. The scenario without any NPIs is set as the baseline. By simulating the epidemic progression and human movement at the individual level, we proposed model-driven targeted interventions, which focus on the surgical testing and quarantine of only a small portion of regions instead of enforcing NPIs in the whole city. The efficacious of common NPIs and the proposed targeted interventions are evaluated by extensive 100 simulations. The proposed model can inform targeted interventions, which are able to effectively contain the COVID-19 outbreak with much lower disruption of the city. It represents a promising approach to sustainable NPIs to help us revive the economy of the city and the world.**

***Index Terms*— Agent-based modeling, Health information management, Local activities, Multilevel systems, Complex networks**

I. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has infected 116,521,281 patients and claimed 2,589,548 lives as of 6:02 pm CET, 8 March 2021 (latest information at: https://covid19.who.int/table). Various nonpharmaceutical interventions (NPIs) were implemented in response to this public health crisis worldwide, ranging from case isolation and quarantine of contacts to the complete lockdown of the entire country [1]. These NPIs were demonstrated to be able to mitigate community transmission at the expense of economic disruption, harm to social and mental well-being, and costly administration required to ensure compliance [2]. However, a resurgence of COVID-19 has been seen in many parts of the world when the NPIs were lifted [3]. As of 8 March 2021, a total of 349,398,520 vaccine doses have been administered, only account for around 4.5% of the global population. The population-wide vaccination program is not likely to cover a large-scale population in developing countries considering the huge demand and difficulty in mass production and supply chains [4], [5]. It is therefore much-needed to find the NPIs to contain outbreaks while minimizing the economic costs [6].

To provide a reliable quantitative analysis of the efficacious of alternative NPIs, we explore a data-driven targeted intervention policy to precisely identify the sub-communities that need to be surgically tested and quarantine. In this study, we develop a data-driven mobility model **(D2M2)** to characterize the demographics, public facilities and functional buildings, transportation system, and travel patterns of the 7.55 million residents in Hong Kong, an international trade center and one of the most densely populated cities in the world. Given the mass amount and rich diversity of the data publicized by the authority and public sectors, Hong Kong represents an ideal opportunity to investigate the feasibility of using a detailed agent-based severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission model to inform targeted interventions. Targeted interventions informed by the model can prioritize a set of high-risk grids for quarantine, so that the localized outbreak can be contained while preserving the normal state in the majority of the city. We examine the efficacy of the proposed targeted interventions in containing the outbreak of COVID-19 in Hong Kong.

Hong Kong was one of the first places outside Mainland China to report the COVID-19 confirmed cases owing to their strong connection [7]. From January to early March 2020, the COVID-19 infected individuals in Hong Kong mainly consisted of imported cases and secondary transmissions, which is considered the first wave of COVID-19 outbreak in Hong Kong. The number of infections in the first wave remained relatively low due to a series of rapid responses, including border control, manual contact tracing, voluntary community-wide use of personal protective equipment (PPE), and social distancing [8], [9]. The second wave mainly consisted of the imported cases outside Asia and associated localized transmission, which occurred from mid-March to May 2020. After implementing more stringent NPIs, only a few local cases were reported in May and early June. On June 18, most social distancing restrictions were lifted. However, the third wave occurred in late June 2020 that pushed Hong Kong's coronavirus case tally to 5,009 by September 19, 2020 [10]. The locally acquired cases appeared on 5 July, which is phylogenetically related to SARS-CoV-2 reported overseas [11]. There could be fourth and more waves to come, as humankind will likely to live with it until a population-wide vaccination program takes place [12]. This study is mainly to simulate the third wave outbreak in Hong Kong from June 25, 2020 to September 24, 2020.

Recently, the Hong Kong government has implemented a compulsory testing policy in response to the epidemic development and the need for infection control [13]. the government will delineate restricted areas and the people within the areas (including the person who had been present in the restricted area for more than two hours in the past 14 days) are required to stay in their premises and undergo compulsory testing and can only leave after the relevant test results are mostly ascertained. The government claims that compulsory

Q. Zhang and D.D. Zeng are the corresponding authors

H. Zhou is with the the school of Data Science, City University of Hong Kong, Hong Kong, China, and the School of Traffic and Transportation Engineering, Central South University (CSU), Changsha, China (e-mail:hanchzhou3-c@my.cityu.edu.hk).

Q. Zhang is with the the school of Data Science, City University of Hong Kong, Hong Kong, China (e-mail:hanchzhou3-c@my.cityu.edu.hk, qingpeng.zhang@cityu.edu.hk).

Z. Cao and D.D. Zeng are with the state Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100091, China (e-mail:zhidong.cao@ia.ac.cn, dajun.zeng@ia.ac.cn).

H. Huang is with the School of Traffic and Transportation Engineering, Central South University (CSU), Changsha, China (e-mail: huanghelai@csu.edu.cn).

testing policy can help to slow down the transmission of SARS-CoV-2 by early identification, early isolation, and early treatment. However, the restricted areas are identified by some index such as the occurrences of new confirmed cases or sewage samples detected COVID-19. This reactive strategy may result in a sudden directive for compulsory testing and suspension of operation of a workplace. The D2M2 can be used to forecast the epidemic dynamic and the targeted restricted areas, which could help to design a more proactive policy and inform the involved residents in advance.

The contribution of this study is threefold. First, the proposed D2M2 is calibrated by the multi-level and multi-source data of the demographics, transportation, and human mobility. To the best of our knowledge, it is the most detailed characterization of the human mobility of Hong Kong. Second, Mobile phone data is difficult to obtain in most countries. Thus D2M2, which is calibrated by open-source data, can be easily extended to the modeling of other metropolises with various demographic and human mobility patterns. Third, based on D2M2, we propose the targeted intervention, which can contain the outbreak with minimal disruption of the society. This is of particular importance to cities like Hong Kong, whose economy relies on international trade.

II. Method

A. Synthetic Population and Their Movement Behavior

To quantitatively characterize the mobility patterns of the residents in Hong Kong, the proposed D2M2 is based on a recent human movement behavior model [14] and is calibrated with the real-world socio-demographic data of Hong Kong [15]. A human-to-human contact and transmission network can be created by combining the human behavior pattern and a synthetic population (as shown in Figure 1). Based on the demographic data from Hong Kong Population Projections and the latest Hong Kong Population Census [16], D2M2 generates synthetic agents with the attributes of age, occupation, ethnicity, and gender. Hence, a total of 7,550,000 agents are generated for Hong Kong residents, of which 5.14 million are individuals between 20 and 65, 9.8 million are students, and the rest are other agents such as elder people and infants. Then each agent is allocated to a corresponding family according to the pre-specific family structure for 2.66 million households. Households were geolocated within 214 discrete areas named Tertiary Planning Units (TPUs) according to spatial characteristics outlined in the census and the TPUs are further split into 4,905 500m $\times$ 500m grids for better aggregated the spatial characteristics of households.

The agents interact in a comprehensive set of Points of Interest (POIs), including houses, workplaces, schools, restaurants, leisure and cultural services venues, shopping malls, pharmacies, groceries, various public entertainment places, etc. The agents interacted at home and community, workplace, school, and other functional buildings. According to the setting previous study [14], during the night, we assume all agents back home and interacted primarily within their households regardless of the weekday or weekends. During the daytime weekday, the agents were divided into three groups: the commuters (i.e., the people who are not working at home) interacted with individuals in the same workplace, and students within their schools while the home-staying agents interact with the individuals who live in the same house or community. In the daytime of weekends, all agents interacted with others according to the spatial distribution of different recreational places (as shown in Figure 2). In addition, the Foreign Domestic Helpers (FDHs) represent an integral part of Hong Kong society nowadays, and 5.29% of the population [17]. We set 39,9320 FDHs among the home working agents according to the statistics of the Immigration Department in 2019 [18]. We set FDHs gather on Sundays in public areas [19], and thus represent the potential risk of localized outbreaks.

B. Transmissiom Model

In the transmission model, infection likelihood between agents is related to the time that he or she spent where infected agents have visited [20].Suppose i and j are two agents. j is infectious, we denote the probability of j infecting i on day t in location type l as:

$$P_{ijl} = \beta \frac{t_l}{t} \tag{1}$$

Here β is the likelihood for individuals to infect one another at specific places and the wider TPU, which is obtained from the basic reproduction number [2], [20]. t_l is the duration of exposure to infection, t is a constant for indoor hours per day. Hence, the probability of individual i getting infected from the location l on day t can be expressed as:

$$P_l(t) = \frac{\beta}{|L|} \sum_{j \in L_{It}} \frac{t_l}{t} \tag{2}$$

where L is a set of individuals of location l, L_{It} is the subset of all individuals who belong to set L and are infectious on day t. $|\cdot|$ denotes the size of the set of individuals and $|L|$ represents the total number of agents in set L. The total number of agents infected in location l on day t $X_L(t)$ can be denoted by a random process:

$$X_L(t) \sim Bin\left(|L| + |L_{It}| + |L_r|, \quad p_l\left(t-1\right)\right) \tag{3}$$

where L_r is the subset of L consisting of agents that have been removed as a result of death or recovery.

C. Model Parameterization and Fitting

A recently introduced susceptible–latent–infected–removed (SLIR) model with some additional compartments is adapted to incorporate the unique characteristics of COVID-19 (as shown in Figure 3) [6], [20], [21]. In particular, at each time step t (one day), the infectious asymptomatic (I_A), infectious symptomatic (I_S) and pre-symptomatic (P_S) individuals can transmit thevirus to susceptible (S) individuals with probability $r\beta$, β and β_s, respectively. If the transmission is successful, the susceptible agent will move to the latent asymptomatic state with probability p or to the latent symptomatic state (L_S) with probability $(1-p)$. A latent asymptomatic individual becomes infectious asymptomatic after a period ϵ'^{-1}, and some latent symptomatic individuals will transit to pre-symptomatic (P_S) state after the same period. Then, during μ^{-1} period after being infected, the symptomatic individuals become onset cases and can either recover or removed after a period μ^{-1}.

The values for the basic reproduction number (R_0) were set as 2.5 according to the literature [22], [23]. Following the setting of the previous study, we ran the models for 80 days to investigate the early stages of a localized outbreak and randomly selected 50 agents as the initial infected on day 0 [20]. The parameters used in the model from literatures are presented in Table I.

To validate the utility of the model set, we simulate the third wave of COVID-19 epidemic in Hong Kong, which began on 25 June 2020 and has been contained in late September after a hard-fought battle. We set the starting date of the simulation to be 25 June 2020 when a cook in a local restaurant had the onset of symptoms, breaking the 21-day streak without local infections. We set the ending date to be 25 September 2020 as the Chief Executive of Hong Kong S.A.R. Government said the third wave of the COVID-19 epidemic has been contained. The infection probability at different locations is

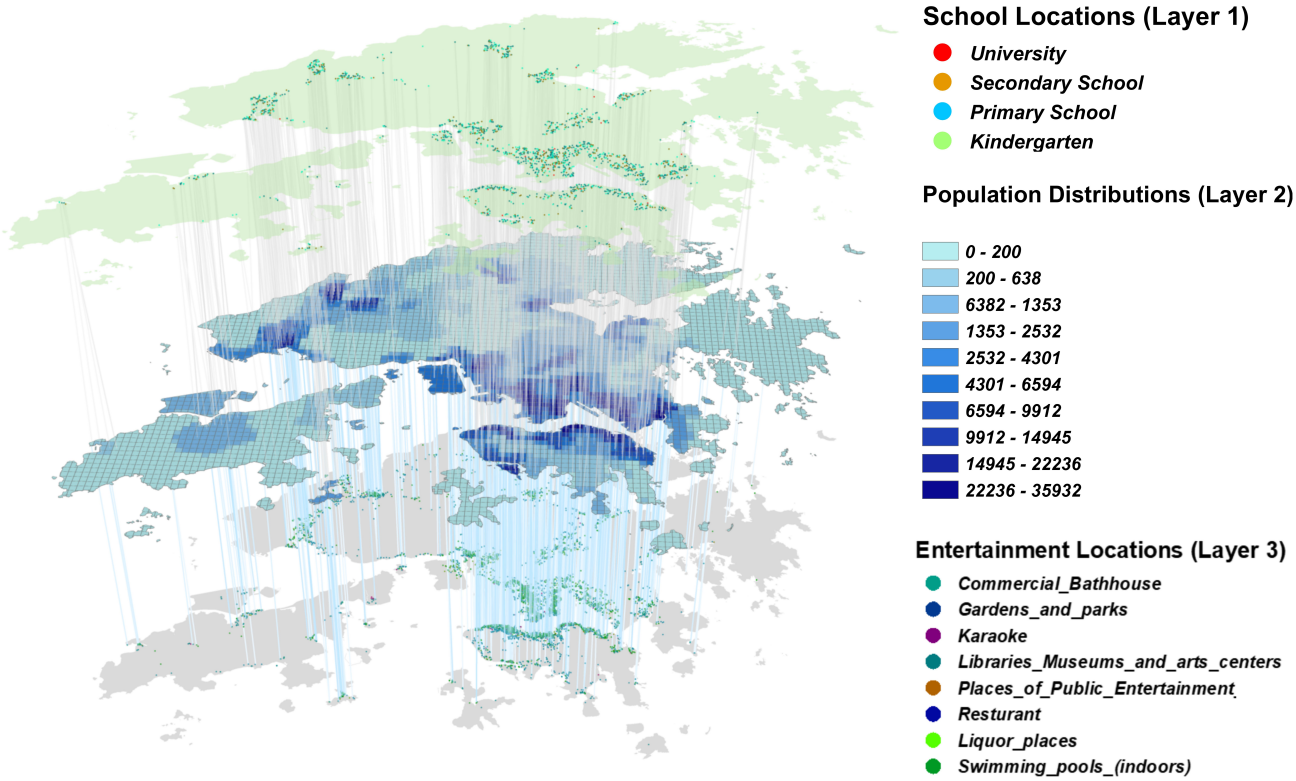


Fig. 1. Geographical distribution of the synthetic population and facilities in Hong Kong. The D2M2 comprises 5.14m adults and 0.92m students and 1.49m other agents (infants or elderly people).

TABLE I
THE PARAMETERS USED IN THE TRANSMISSION MODEL

Parameters	Description	Value	References
γ	Relative infectiousness of asymptomatic individuals	50%	[24]
ϵ^{-1}	Incubation period	6.2 days	[25]
ϵ'^{-1}	Latent period (for asymptomatic and symptomatic case)	2.9 days	[21]
pre	pre-symptomatic period	ϵ^{-1}-ϵ'^{-1} days	[21]
p	Proportion of asymptomatic	17.7%	[26]
μ^{-1}	Time to removed/recovery	2.9 days	[23]
k	Proportion of presymptomatic transmission	12.6%	[27]
R_0	Basic reproduction number	2.5	[6]
β	Basic transmission rate	$\frac{R_0\mu}{pr+(1-p)\div(1-k)}$	[6]

being updated based on the real-time responses of the government. After the Education Bureau announced *school closure* on 13 July [28], we assume the infection probability at school to be 0 and the residents' infection probability at other locations are being updated according to Google's Community Mobility Reports. On 19 July, the Hong Kong S.A.R. Government announced a series of anti-pandemic measures such as the bans on dine-in services at restaurants from 6 pm to 5 am (*anti-pandemic measures*). On 27 July, the government further tightened social distancing measures, including gathering in public will be limited to only two people per group and restaurants are disallowed to offer dine-in services (*tightened measures*). These measures may result in a huge decrease in contact intensity. Thus, based on the previous studies and surveys [29], we assumed that on 19 July, the contact intensity decreased to 50% and then further decrease to 10% on 27 July of the normal situation at the workplace and other public places respectively. Moreover, we also assumed the people who live or work with someone diagnosed with COVID-19 will follow the Home Quarantine Rule [30]. These agents will complete the 14 full days of self-isolation (separate themselves from others at home) and cannot infect other agents.

D. Targeted Interventions, Social-distancing, and Testing Strategies

To confront the pandemic with less disruption, we propose the *targeted interventions*, which identify a small subset of areas for quarantine while maintaining mobility in other areas. Specifically, we run 200 simulations of *baseline scenario* (i.e., no NPIs) with 50 randomly seeded agents. Then,following Koo's study [20], based on the outcome of the baseline scenario simulations at day 80, thegrids are ranked by the number of expected infections in the baseline simulation. The top 50 and top 100 grids are chosen as the High-Risk Grids (HRGs), while the others are non-HRGs. Compulsory quarantine is taking place at HRGs. Within HRGs, all essential materials are uniformly delivered by the authorities. Residents in non-HRGs are not affected except that they could not freely commute to HRGs. As a comparison, we examine the efficacy of the *stay-at-home control*, which involves the following scenario starting from day 15 (the actual delay in the response in Hong Kong [31]): (a) School closures. We transform all school students to home working agents simultaneously. (b) Social distancing. We assume that all places are still open, but the volume of passengers is proportional to the Google Mobility Report on the same day.

Moreover, we examine the efficacy of the *reactive control*. Testing is a useful tool to detect, predict, and reduce the spread of disease [32], [33]. In this scenario, in addition to the *stay-at-home* control, patients with symptoms are quickly tested and quarantined since the

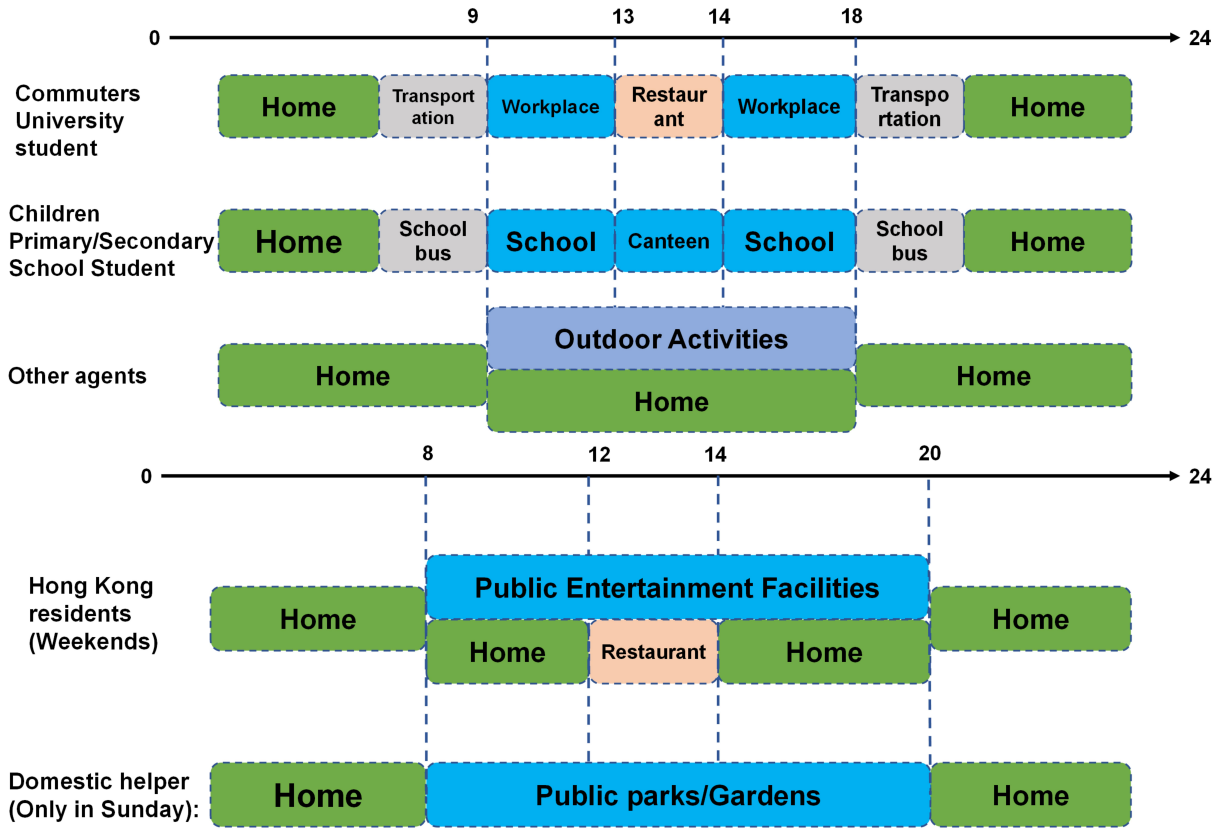


Fig. 2. Schematic illustration of the daily movement of three groups of people in Hong Kong on weekdays and weekends. Commuters include people aged 23–60 who work outside home or study at universities. Students includes children and adolescents aged 4–18 who are studying at schools (excluding university); Others includes children less than 4 years of age, retired elderly adults (> 65), the unemployed, and those who work from home.

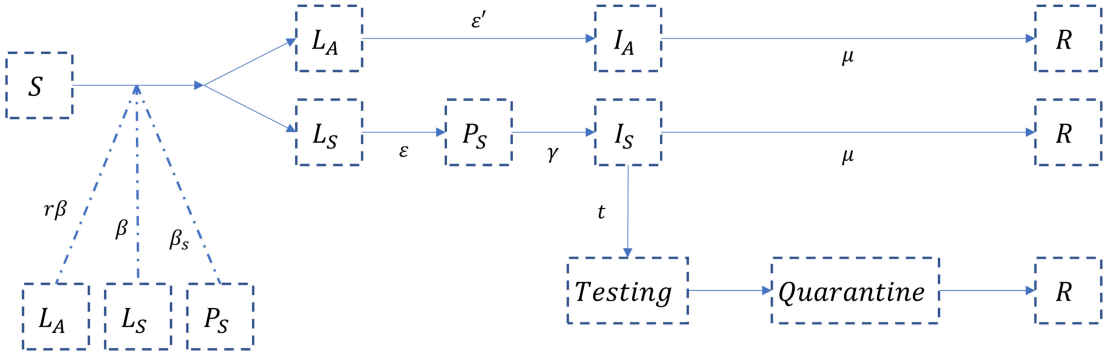


Fig. 3. The SLIR model. Here we denote susceptible (S), latent asymptomatic (L_A), latent symptomatic (L_S), presymptomatic (P_S), infectious asymptomatic (I_A), infectious symptomatic (I_S), testing symptomatic case ($Testing$), positive in testing and quarantine ($Quarantine$) and recovered/removed (R) individuals.

authorities respond on day 15 [34]. We set that testing results will come out either 24 or 48 after being tested [6]. In the experiment, we evaluate the outcomes with different proportions of symptomatic patients who get tested(10%, 20%, 50%). The maximum testing capacity is set as 10,000 cases per day according to the governments' publication [35]. Finally, we investigate the *full lockdown control* on the top of testing 50% of symptomatic cases, which represents the scenario that school closer the 70% reduction in human mobility is taken place on day 1, which follows the setting of strict lockdown in the study of US National Bureau of Economic Research [36]. Note that we run 100 simulations for non-baseline NPIs (summarized in Table II) to account for the stochasticity and to calculate the confidence intervals (CIs).

III. Results

A. Compare the Simulation with Real-world Data

Figure 4 presents the comparison of the simulation results of the proposed D2M2 model with a random number of initial infections (5 to 15) and the number of confirmed cases in Hong Kong during the third wave. Note that we use the onset date instead of the reported date of each infected case and exclude the asymptomatic cases to reflect the real situation. In general, the simulation results (blue curve) capture the trend and turning point of the confirmed case data (grey bars). The real-world data are within the interquartile range (IQR) of

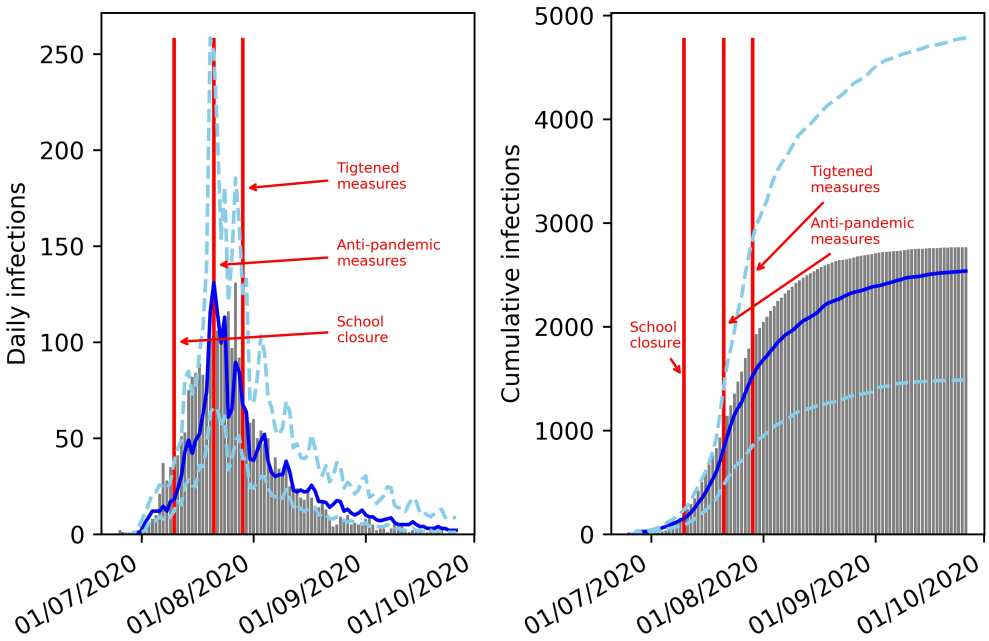


Fig. 4. Comparison of simulated results with the official confirmed case. The total number of daily infections is shown on the left; the cumulative number of infections is shown on the right. The blue curve (dashed curve for IQR) and grey bars represent the simulated results and the confirmed cases, respectively. The red vertical lines indicate the actual control measures in Hong Kong.

TABLE II
SUMMARY OF APPLIED NPIS

NPIs	Description
Baseline	No intervention
Stay-at-home control	School closure + Mobility changed
Reactive control	*Stay-at-home control* + testing 10%, 20% or 50% of symptomatic patients
Full lockdown control	*Strictest reactive control* + 70% mobility reduction
Targeted interventions	Controlling 50 or 100 girds facing the highest infection risk

the model simulations. The blue curve (the median of simulations) is a little lower than the real-world data because of the superspreading events, which are likely to result in more explosive growth [37]. Specifically, our model is a theoretical model that simulates viral transmission based on the general population movement pattern and average disease transmissibility, and thus lack of data from large clusters or superspreading events to simulate the real-world situations [38]. The implementations of the control measures are represented by the red lines. The effects of anti-pandemic and further tightened measures became obvious with fluctuations, due to the incubation period and asymptomatic transmission. Moreover, the effect of school closure just slightly affects the infection number.

B. Quantify the Effectiveness of NPIs

For the baseline scenario, a cumulative median of around 129,000 cases (IQR: 94,000-193,000) of the Hong Kong population was infected on day 80 (Figure 3, Table III). At this level of transmission (R_0=2.5), mild social distancing resulted in a median of around 17,000 cases (IQR: 12,000–25,000), the full lockdown (city-wide quarantine) measure led to 805 cases (IQR: 636–10,60). Regrading reactive control measure with a 24-hour-delay, testing 10% of onset cases led to in a median of around 13,000 cases (IQR: 9,000 - 21,000); testing 20% of onset cases led to 5,830 cases (IQR: 4,315-9,204); Testing 50% led to 2,667 cases (IQR: 2,187-3,618). With a 48-hour-delay, the infected cases number resulted by testing 10%, 20%, and 50% symptomatic cases would be 16,000 (IQR: 9,000-31,000), 9,761 (IQR: 7,584-17,675) and 5,649 (IQR: 3,691-9,433) respectively. Compared with the baseline scenario, the full lockdown led to the greatest reduction in cases (99.32%, IQR: 95.83%-99.37%).

The location of the infection (school, workplace, domestic party, non-commuter place or home community) was set as the site where individuals acquired the virus from others. As shown in Table 3, in most scenarios, the home was the place with the highest number of infections. Workplace contributed the second-highest number of infectious, indicating that reducing the mobility in the workplace could reduce the risk of infections since the commuters interacted with agents in the same workplaces or functional buildings [20]. School closure could reduce the infectious among students but contributed little to protect other residents. This has also been confirmed by the previous study [29]. Moreover, the number of infections at FDHs' gathering of other POIs was relatively consistent and count small proportions in different NPIs. In the full lockdown scenario, the infections mainly occurred at the POIs that are neither school or workplace, including grocery, pharmacy, or retail store. This indicates we may further reduce the number of infections by reducing the activities in POIs at the cost of wider social disruption.

The epidemic peak has not been reached by day 80 in all scenarios, because here we did not consider the contact tracing for model simplicity. As shown in Figure 3, the median number of newly infected cases in the baseline and mild social distancing scenarios was consistently increasing, whereas the number of new cases was relatively stable under the instantaneous response. From the spatial perspective, a considerable number of new infections were acquired from the urban areas in Kowloon and New Territories (Figure 5). The instantaneous response can effectively reduce the spillover effect of the epidemic outbreak from urban to suburban area. The infection distributions in the baseline and mild social distancing scenarios were

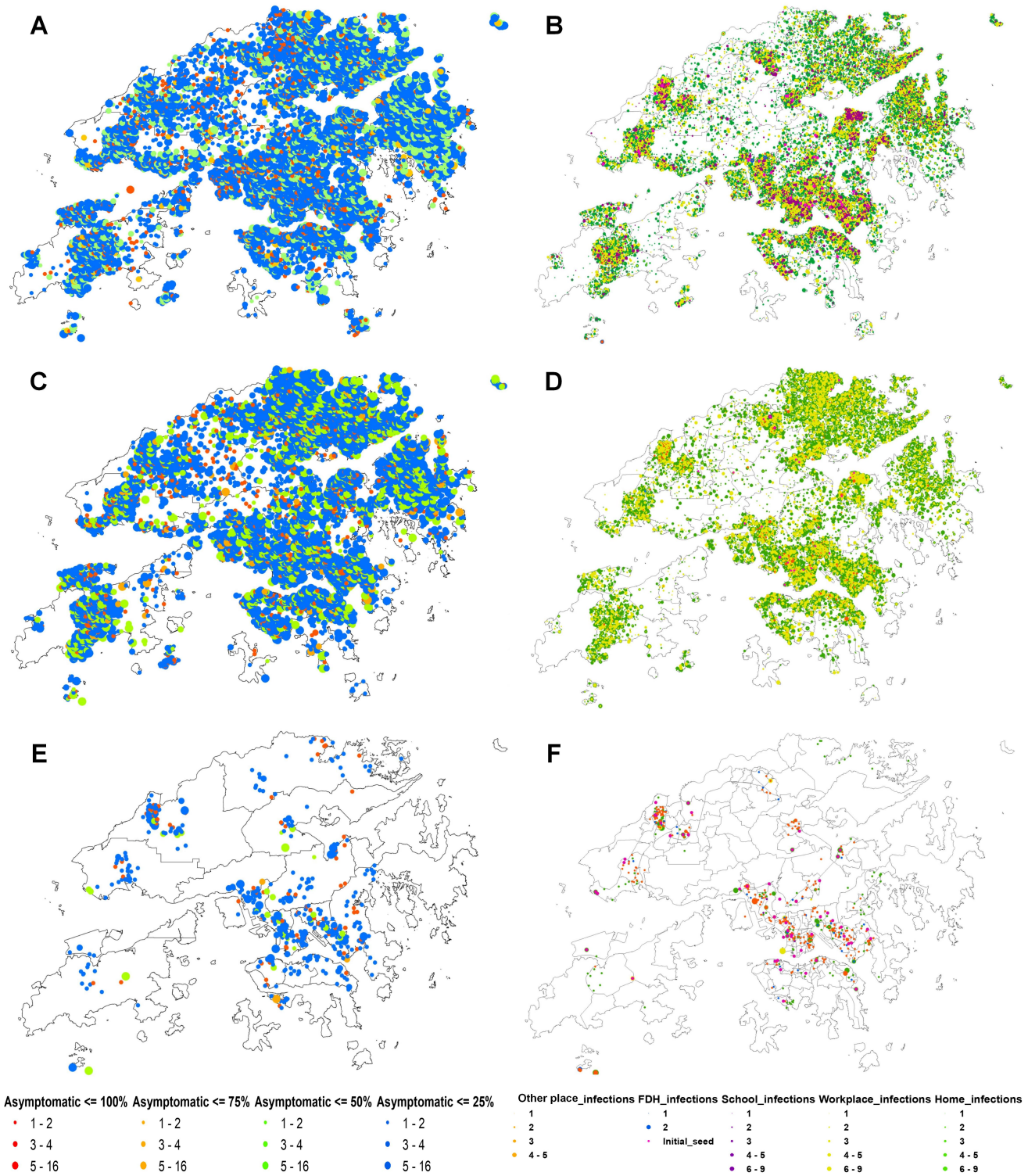


Fig. 5. Geospatial distribution of the simulated infections on day 80. The locations are determined by the household location of infected individuals. The size of dots is proportional to the number of infections in the corresponding household. For panels A, C, and E, the color of dots represents the percentage of asymptomatic cases in each household. Blue: 0%-25%; Green: 25%-50%; Orange: 50%-75%; Red: 75%-100%. For panel B, D, and F, the color of dots represents the source of the infection. Green: household; Yellow: workplace; Purple: school; Blue: gathering of domestic helpers; Orange: Other POIs.

more widely distributed in less population-dense areas. Given the dense residential clusters mainly sited on the Kowloon, Hong Kong island, and the new towns of New Territories, the Large-scale and homogeneously distributed infections implied that the virus cannot be contained without massive testing or total lockdown strategy.

C. Results of Targeted Interventions

The targeted interventions were derived as follows. First, we ran simulations of the baseline scenario (Figure 7). Second, the top 100 and 50 grids with the highest number of infections were chosen as the HRGs (Figure 8), which will be quarantined, whereas the people who

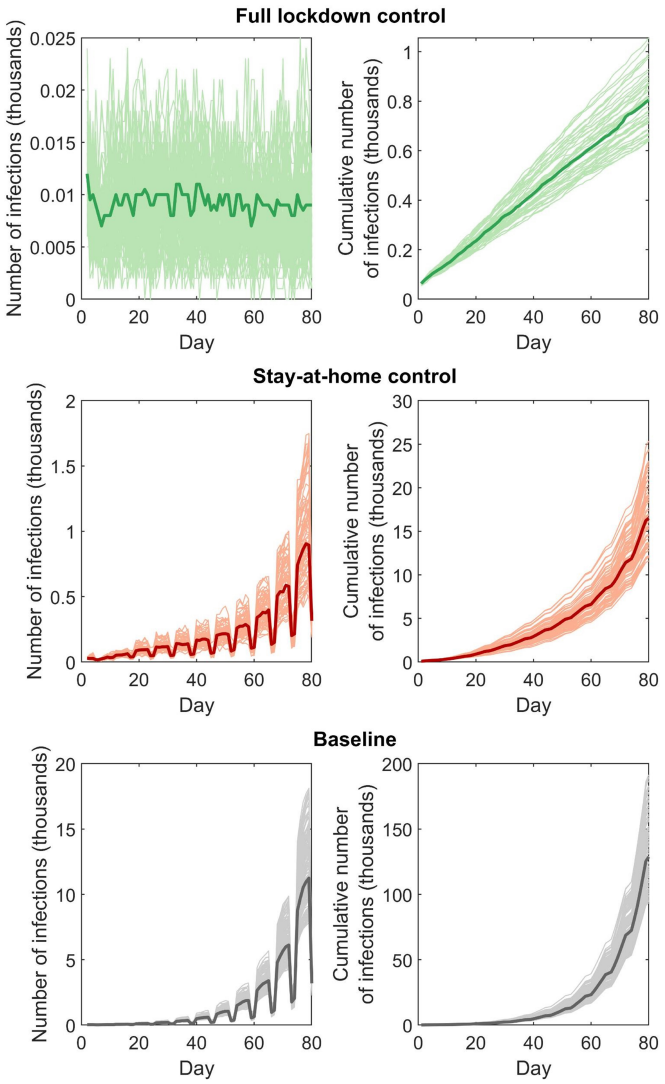


Fig. 6. Total daily number and cumulative number of SARS-CoV-2 infections up to 80 days after different control measures Total number of daily infections is shown on the left; cumulative number of infections is shown on the right. Dark lines represent the medians in each panel. Shaded areas represent the MC simulation results.

live in other non-HRG grids were not affected. Therefore, the targeted interventions are proactive NPIs, which are based on the prediction of the epidemic patterns. Simulation results showed that the targeted interventions could effectively reduce the size of the outbreak in the city, with 5,342 (IQR 4,578-6,607) and 7,550 (6,660-8,781) accumulative infected individuals after 80 days by only quarantining 100 (2.04% of the area and 24.46% of the population) or 50 (1.02% of the area and 13.05% of the population) grids in the city. This indicated that the COVID-19 outbreak was largely contained without much disruption of society. The efficacy of quarantining 100 HRGs slightly outperformed the reactive control strategy with 50% symptomatic cases tested within 48 hours (Figure 9). Considering the high cost of contact tracing and mass testing, the targeted interventions stand a good potential to become reliable and cost-efficient NPIs to confront the long-term COVID-19 pandemic and even endemic.

Geospatially, the top 50 HRGs are mainly located in urban areas in Kowloon and New Territories, two of the most densely populated districts in the world. More grids in Hong Kong Island were identified as in the top 50-100 HRGs. The targeted interventions had a

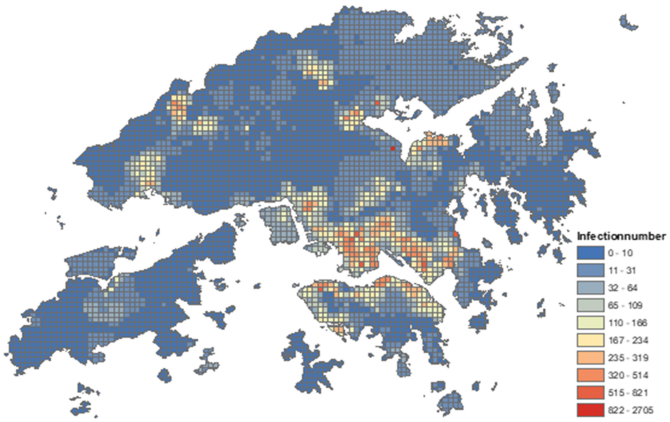


Fig. 7. The heat map for the infections across the city in the baseline scenario (no NPIs).

suppressive effect on new infections over time, though the number of cases continued to increase, and the cumulative number of cases are exponentially growing, implying a delayed intervention effect rather than a preventive effect here as observed in the full lockdown scenario.

D. Results of Reactive Control Measure

We examined the effect of testing by simulating three scenarios that 10%, 20%, and 50% percent of symptomatic cases are tested after the first 15 days (Figure. 6). Moreover, to analyze the effect of testing speed, the testing results are set to come out 24-hour or 48-hour after the testing. Compared with the baseline data, testing 10%, 20%, 50% of symptomatic cases within 24 hours reduced the infectious by 89.92% (IQR 88.88%-90.64%), 95.47% (95.23%-95.87%) and 97.93% (97.66%-98.13%), respectively. As contrast, testing 10%, 20%, 50% of symptomatic cases within 48 hours reduced in 87.78% (83.85% - 90.44%), 92.42% (91.89%-92.84%) and 95.61% (95.11% - 96.05%), respectively. Apparently, the speed of testing greatly affects the efficacy of NPIs. These results indicate that timely and prompt testing is essential to such reactive control measure [39].

Regarding the location of infections, the home-community and workplace ranked 1st and 2nd in all testing scenarios. More importantly, due to the stay-at-home and testing strategies, the home-community was responsible for more than half of infections (Table III). In the scenario of 24-hour testing delay, the home-community and workplace infections accounted for 89.76% (88.73%-90.81%), 88.78% (83.84%-88.93%) and 79.79% (74.89%-81.14%) at 10%, 20% and 50% testing scale level, respectively. The corresponding values were 88.62% (88.09%-88.93%), 89.04% (87.31%-90.30%) and 88.52% (80.89%-88.66%) in the scenario of 48-hour testing delay. Compared to the full lockdown scenario, the wide gap between the infection sizes implied that the community and workplace-based transmission is dramatically growing and became hard to control if no strict movement restrictions were implemented during the outbreak period.

IV. Discussion

COVID-19 has resulted in enormous damage to the society and economy. Many countries have been suffering from a shortage of testing resources and economic stagnation due to the necessary lockdowns. The quantitative analyses with D2M2, which is calibrated with the real-world big data of human mobility, demonstrate that

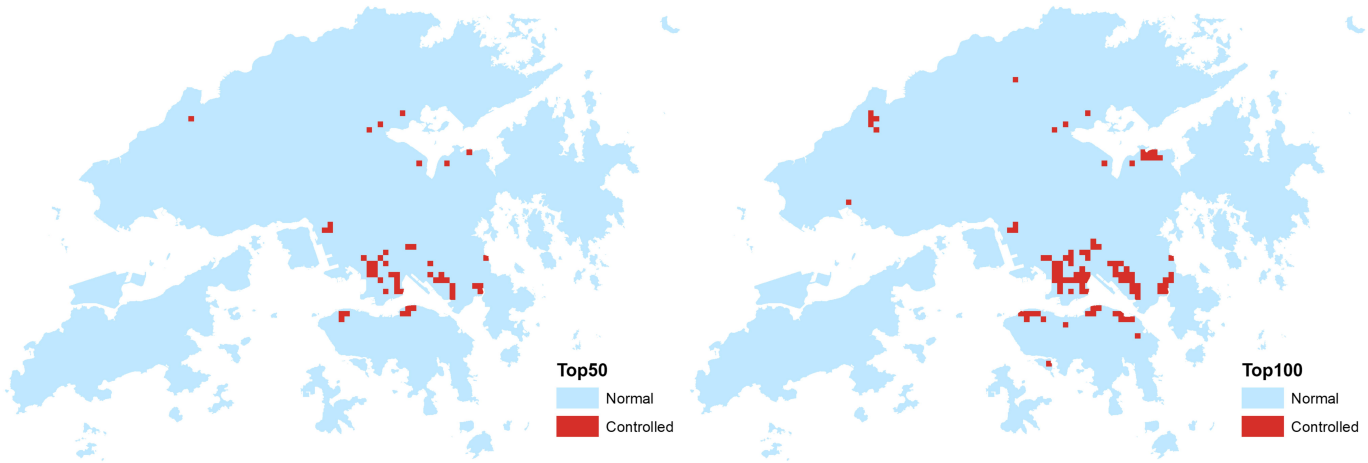


Fig. 8. The geospatial distributions of identified HRGs (left: top 50 grids, right: top 100 grids).

TABLE III
ESTIMATED MEDIAN (IQR) OR CUMULATIVE NUMBER OF SARS-COV-2 INFECTIONS ON DAY 80 BY LOCATION, INTERVENTION, AND TESTING RESULTS DELAY

	Baseline	Stay at home control	Testing 10% onset infections	Testing 20% onset infections	Testing 50% onset infections	Full lockdown control
	No testing scenario		**One-day (24h) turn around time for testing result**			
Total number of infections	128711 (93530-192972)	16503 (11610-25426)	12969 (8756-21462)	5830 (4315-9204)	2667 (2187-3618)	805 (636-1060)
Home community	68681 (49049-109763)	10642 (7467-16577)	8427 (5632-14326)	3787 (2613-5977)	1582 (1205-2218)	286 (222-389)
School	26115 (18421-40856)	84 (70-97)	41 (35-53)	42 (35-52)	41 (35-47)	15 (11-20)
Workplace	31103 (24657-42614)	4440 (3100-6581)	3215 (2137-5163)	1389 (1005-2173)	546 (433-718)	103 (75-131)
FHDs gathering	15 (6-42)	12 (4-30)	13 (5-39)	15 (6-40)	10 (4-26)	10 (3-23)
Other POIs	571 (453-760)	444 (365-606)	373 (248-524)	320 (228-512)	339 (232-445)	404 (267-522)
	Controlled 50 HRGs	Controlled 100 HRGs	Testing 10% onset infections	Testing 20% onset infections	Testing 50% onset infections	-
	Targeted interventions		**Two-day (48h) turn around time for testing result**			
Total number of infections	7550 (6660-8781)	5342 (4578-6607)	15730 (8940-31158)	9761 (7584-17675)	5649 (3691-9433)	- -
Home community	4029 (3520-4731)	2829 (2424-3517)	10487 (5866-21594)	6250 (4883-11941)	3777 (2214-6346)	- -
School	1095 (971-1234)	681 (572-806)	44 (35-54)	41 (38-52)	39 (33-47)	- -
Workplace	1420 (1236-1643)	941 (811-1175)	3307 (2057-6117)	2272 (1870-4020)	1224 (772-1999)	- -
FHDs gathering	14 (5-38)	16 (6-42)	13 (3-31)	10 (4-33)	12 (5-42)	- -
Other POIs	576 (448-807)	578 (457-812)	345 (256-498)	392 (291-562)	290 (218-422)	- -

All numbers have been rounded to the nearest integer and there will be some discrepancies in the summations

targeted interventions are feasible NPIs to contain the outbreak with only a small fraction of areas and population being affected.

Apparently, the most effective strategy to contain the COVID-19 outbreak is to have a stringent lockdown, which is also echoed by our results. However, city-wide quarantine and lockdown may lead to additional loss including the elevating unemployment rate, social isolation, and mental health crisis [40], [41]. Moreover, COVID-19 has the potential to become an endemic [12], if long-term protection of vaccines is not satisfactory. Therefore, many countries would have to lift the NPIs with a certain level of COVID-19 risk for the near future [42]. In practice, the proposed targeted intervention strategy can partially mitigate isolation fatigue and social disruption by proactively identifying only a small subset of the area and population for testing and quarantine, while preserving the mobility of most residents. Furthermore, as an agent-based model, D2M2 is versatile and can be modified to incorporate other factors such as vaccination programs.

In practice, as an international financial center and a major supply chain hub, Hong Kong is likely to face the continuous risk of COVID-19 in the near future. The proposed D2M2 and associated targeted interventions have the potential to guide the current compulsory COVID-19 testing policy achieves a balance between lowering the risk and preserving the human mobility and economy of the city. This also applies to other major cities in the world, such as Beijing, New York, London, Tokyo, etc. In the big data era, it is essential and promising to capitalize on the well of valuable data about human mobility and urban structure to benefit mankind in various contexts. Our study is an effort in combining real-world data of population

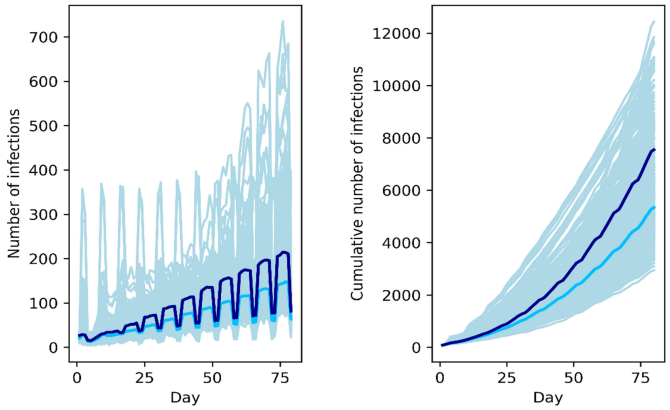


Fig. 9. Results of the targeted interventions. The total number of daily infections is shown on the left; cumulative number of infections is shown on the right. The dark navy lines represent the corresponding median when 50 grids were controlled. Light blue lines represent the medians in each panel when 100 grids were controlled. Light blue shading represents all 200 simulations for each scenario.

demographics, public facilities and functional buildings, transportation systems, and travel patterns to inform effective NPIs to combat emerging infectious diseases like COVID-19. The limitation of this study is the missing of real-time individualized mobility movement data such as the mobile phone data. Obtaining individualized mobile phone data is extremely difficult, if not impossible. However, the model can be applied to other cities where such individualized mobility data is available.

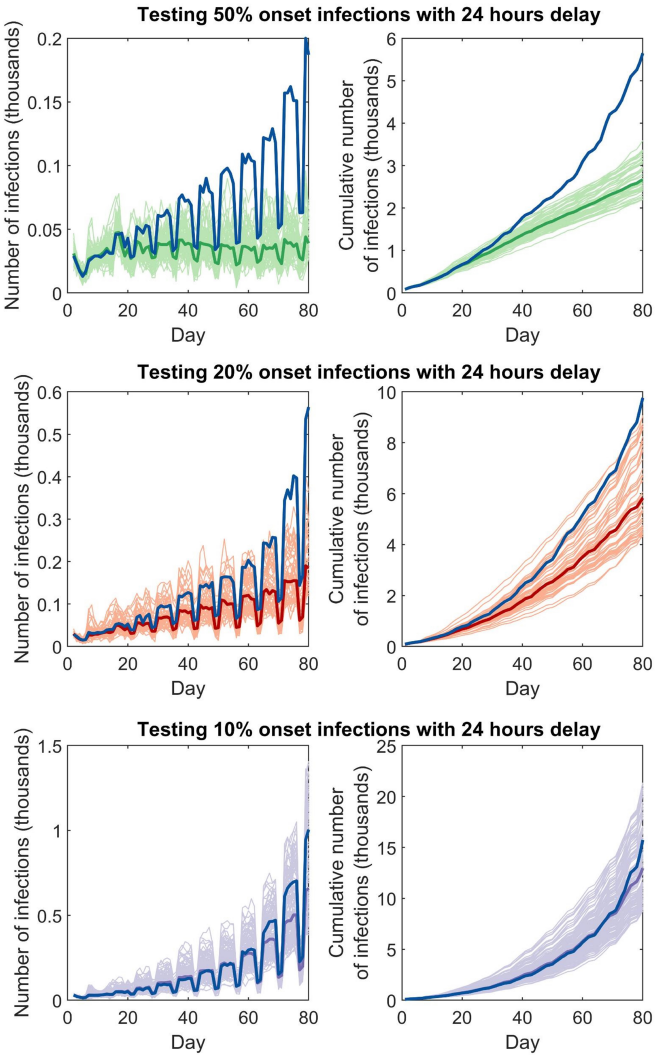


Fig. 10. Total daily number and cumulative number of SARS-CoV-2 infections up to 80 days after different 24 hours delay testing scenarios. Total number of daily infections is shown on the left; cumulative number of infections is shown on the right. Dark green, dark orange and purple lines represent the median number when there 24 hours delay for tests. Dark navy lines represent the corresponding medians of 48 hours output delay as a comparison. Shaded areas show all 100 simulations for each scenario.

References

[1] S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton *et al.*, "Estimating the effects of non-pharmaceutical interventions on covid-19 in europe," *Nature*, vol. 584, no. 7820, pp. 257–261, 2020.

[2] B. L. Dickens, J. R. Koo, J. T. Lim, M. Park, S. Quaye, H. Sun, Y. Sun, R. Pung, A. Wilder-Smith, L. Y. A. Chai *et al.*, "Modelling lockdown and exit strategies for covid-19 in singapore," *The Lancet Regional Health-Western Pacific*, vol. 1, p. 100004, 2020.

[3] P. Tupper, H. Boury, M. Yerlanov, and C. Colijn, "Event-specific interventions to minimize covid-19 transmission," *Proceedings of the National Academy of Sciences*, vol. 117, no. 50, pp. 32 038–32 045, 2020.

[4] P. G. Walker, C. Whittaker, O. J. Watson, M. Baguelin, P. Winskill, A. Hamlet, B. A. Djafaara, Z. Cucunubá, D. O. Mesa, W. Green *et al.*, "The impact of covid-19 and strategies for mitigation and suppression in low-and middle-income countries," *Science*, vol. 369, no. 6502, pp. 413–422, 2020.

[5] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, and M. Lipsitch, "Projecting the transmission dynamics of sars-cov-2 through the post-pandemic period," *Science*, vol. 368, no. 6493, pp. 860–868, 2020.

[6] A. Aleta, D. Martin-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini Jr, S. Merler *et al.*, "Modelling the impact of testing, contact tracing and household quarantine on second waves of covid-19," *Nature Human Behaviour*, vol. 4, no. 9, pp. 964–971, 2020.

[7] B. J. Cowling, S. T. Ali, T. W. Ng, T. K. Tsang, J. C. Li, M. W. Fong, Q. Liao, M. Y. Kwan, S. L. Lee, S. S. Chiu *et al.*, "Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in hong kong: an observational study," *The Lancet Public Health*, vol. 5, no. 5, pp. e279–e288, 2020.

[8] H. Y. Lam, T. S. Lam, C. H. Wong, W. H. Lam, C. M. E. Leung, K. W. A. Au, C. K. Y. Lam, T. W. W. Lau, Y. W. D. Chan, K. H. Wong *et al.*, "The epidemiology of covid-19 cases and the successful containment strategy in hong kong–january to may 2020," *International Journal of Infectious Diseases*, vol. 98, pp. 51–58, 2020.

[9] V. C.-C. Cheng, S.-C. Wong, V. W.-M. Chuang, S. Y.-C. So, J. H.-K. Chen, S. Sridhar, K. K.-W. To, J. F.-W. Chan, I. F.-N. Hung, P.-L. Ho *et al.*, "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (covid-19) epidemic due to sars-cov-2," *Journal of Infection*, vol. 81, no. 1, pp. 107–114, 2020.

[10] T. government of Hong Kong Special Administrative Region. Together, we fight the virus. [Online]. Available: https://www.coronavirus.gov.hk/sim/index.html

[11] K. K.-W. To, W.-M. Chan, J. D. Ip, A. W.-H. Chu, A. R. Tam, R. Liu, A. K.-L. Wu, K.-C. Lung, O. T.-Y. Tsang, D. P.-L. Lau *et al.*, "Unique clusters of severe acute respiratory syndrome coronavirus 2 causing a large coronavirus disease 2019 outbreak in hong kong," *Clinical Infectious Diseases*, 2020.

[12] J. Shaman and M. Galanti, "Will sars-cov-2 become endemic?" *Science*, vol. 370, no. 6516, pp. 527–529, 2020.

[13] C. E. in Council. (2020, Nov) Prevention and control of disease (compulsory testing for certain persons) regulation. [Online]. Available: https://www.elegislation.gov.hk/hk/cap599J

[14] N. Zhang, H. Huang, B. Su, X. Ma, and Y. Li, "A human behavior integrated hierarchical model of airborne disease transmission in a large city," *Building and environment*, vol. 127, pp. 211–220, 2018.

[15] N. Zhang, W. Jia, H. Lei, P. Wang, P. Zhao, Y. Guo, C.-H. Dung,

Z. Bu, P. Xue, J. Xie *et al.*, "Effects of human behaviour changes during the covid-19 pandemic on influenza spread in hong kong," *Clinical Infectious Diseases*, 2020.

[16] Census and S. D. of the government of Hong Kong Special Administrative Region. (2017, Dec) Population by-census 2016. [Online]. Available: https://www.bycensus2016.gov.hk/en/bc-dp-tpu.html

[17] I. D. of the government of Hong Kong Special Administrative Region. Statistics on the number of foreign domestic helpers in hong kong. [Online]. Available: https://data.gov.hk/en-data/dataset/hk-immd-set4-statistics-fdh

[18] L. C. Commission *et al.*, "Foreign domestic helpers and evolving care duties in hong kong," *Hong Kong: Research Office, Legislative Council Secretariat*, 2017.

[19] C. B. Patterson, "Queer, brown, migrant: documenting the hong kong 'helper'," *Cultural Studies*, vol. 33, no. 6, pp. 1008–1028, 2019.

[20] J. R. Koo, A. R. Cook, M. Park, Y. Sun, H. Sun, J. T. Lim, C. Tam, and B. L. Dickens, "Interventions to mitigate early spread of sars-cov-2 in singapore: a modelling study," *The Lancet Infectious Diseases*, vol. 20, no. 6, pp. 678–688, 2020.

[21] X. Hao, S. Cheng, D. Wu, T. Wu, X. Lin, and C. Wang, "Reconstruction of the full transmission dynamics of covid-19 in wuhan," *Nature*, vol. 584, no. 7821, pp. 420–424, 2020.

[22] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study," *The Lancet*, vol. 395, no. 10225, pp. 689–697, 2020.

[23] S. M. Moghadas, A. Shoukat, M. C. Fitzpatrick, C. R. Wells, P. Sah, A. Pandey, J. D. Sachs, Z. Wang, L. A. Meyers, B. H. Singer *et al.*, "Projecting hospital utilization during the covid-19 outbreaks in the united states," *Proceedings of the National Academy of Sciences*, vol. 117, no. 16, pp. 9122–9126, 2020.

[24] H. Nishiura, T. Kobayashi, T. Miyama, A. Suzuki, S.-m. Jung, K. Hayashi, R. Kinoshita, Y. Yang, B. Yuan, A. R. Akhmetzhanov *et al.*, "Estimation of the asymptomatic ratio of novel coronavirus infections (covid-19)," *International journal of infectious diseases*, vol. 94, p. 154, 2020.

[25] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, "The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application," *Annals of internal medicine*, vol. 172, no. 9, pp. 577–582, 2020.

[26] H. D. of the government of Hong Kong Special Administrative Region. Statistics on the number of foreign domestic helpers in hong kong. [Online]. Available: https://www.chp.gov.hk/files/pdf/local_situation_covid19_en.pdf

[27] Z. Du, X. Xu, Y. Wu, L. Wang, B. J. Cowling, and L. A. Meyers, "Serial interval of covid-19 among publicly reported confirmed cases," *Emerging infectious diseases*, vol. 26, no. 6, p. 1341, 2020.

[28] E. B. of the government of Hong Kong Special Administrative Region. Arrangements of early commencement of summer holiday for all schools. [Online]. Available: https://www.edb.gov.hk/attachment/en/sch-admin/admin/about-sch/diseases-prevention/edb_20200710_eng.pdf

[29] J. Yang, Q. Zhang, Z. Cao, J. Gao, D. Pfeiffer, L. Zhong, and D. D. Zeng, "The impact of non-pharmaceutical interventions on the prevention and control of covid-19 in new york city," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 31, no. 2, p. 021101, 2021.

[30] S. F. D. of Public Health. Home isolation and quarantine guidelines for people with coronavirus-2019 infection and their household or close contacts. [Online]. Available: https://www.sfcdcp.org/wp-content/uploads/2020/04/COVID19-Home-IQ-Guidelines-FINAL-05.14.2020.pdf

[31] D. Wei. Hong kong shuts down schools again as virus cases surge. [Online]. Available: https://www.bloomberg.com/news/articles/2021-04-24/bill-hwang-s-implosion-bruises-billionaire-teacher-from-china

[32] W. E. Allen, H. Altae-Tran, J. Briggs, X. Jin, G. McGee, A. Shi, R. Raghavan, M. Kamariza, N. Nova, A. Pereta *et al.*, "Population-scale longitudinal mapping of covid-19 symptoms, behaviour and testing," *Nature Human Behaviour*, vol. 4, no. 9, pp. 972–982, 2020.

[33] Y. Xing, G. W. Wong, W. Ni, X. Hu, and Q. Xing, "Rapid response to an outbreak in qingdao, china," *New England Journal of Medicine*, vol. 383, no. 23, p. e129, 2020.

[34] V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S. Brünink, J. Schneider, M. L. Schmidt *et al.*, "Detection of 2019 novel coronavirus (2019-ncov) by real-time rt-pcr," *Eurosurveillance*, vol. 25, no. 3, p. 2000045, 2020.

[35] D. of Health and H. A. of the government of Hong Kong Special Administrative Region. Statistics on testing for covid-19 in hong kong. [Online]. Available: https://www.chp.gov.hk/files/pdf/statistics_on_covid_19_testing.pdf

[36] P. Fajgelbaum, A. Khandelwal, W. Kim, C. Mantovani, and E. Schaal, "Optimal lockdown in a commuting network," National Bureau of Economic Research, Tech. Rep., 2020.

[37] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, "Superspreading and the effect of individual variation on disease emergence," *Nature*, vol. 438, no. 7066, pp. 355–359, 2005.

[38] K. Leung, J. T. Wu, and G. M. Leung, "Real-time tracking and prediction of covid-19 infection using digital proxies of population mobility and mixing," *Nature communications*, vol. 12, no. 1, pp. 1–8, 2021.

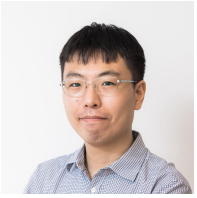
[39] D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, and R. Parker, "Test sensitivity is secondary to frequency and turnaround time for covid-19 screening," *Science advances*, vol. 7, no. 1, p. eabd5393, 2021.

[40] T. Lancet, "India under covid-19 lockdown," *Lancet (London, England)*, vol. 395, no. 10233, p. 1315, 2020.

[41] T. J. VanderWeele, "Challenges estimating total lives lost in covid-19 decisions: consideration of mortality related to unemployment, social isolation, and depression," *Jama*, vol. 324, no. 5, pp. 445–446, 2020.

[42] Y. Li, H. Campbell, D. Kulkarni, A. Harpur, M. Nundy, X. Wang, H. Nair, U. N. for COVID *et al.*, "The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (r) of sars-cov-2: a modelling study across 131 countries," *The Lancet Infectious Diseases*, vol. 21, no. 2, pp. 193–202, 2021.

First A. Author received the Bachelor's degree in Traffic Engineering and Software Engineering from Dalian Jiaotong University, Liaoning, China, in 2014. And received the master's degree in Communication and Transportation Engineering from Central South University, Hunan, China, in 2017. He is currently working toward a joint PhD degree with Central South University, Hunan, China and City University of Hong Kong, Hong Kong, China. His research area includes Data analysis,intelligent transportation systems, and statistical analysis



Qingpeng Zhang received the B.S. degree in automation from the Huazhong University of Science and Technology, Wuhan, China, in 2009, and the Ph.D. degree in systems and industrial engineering from the University of Arizona, Tucson, AZ, USA, in 2012.

He was a Post-Doctoral Research Associate with the Department of Computer Science, Tetherless World Constellation, Rensselaer Polytechnic Institute, Troy, NY, USA. He is currently an Associate Professor with the School of Data Science, City University of Hong Kong, Hong Kong. His current research interests include medical informatics and network science. Dr.Zhang is an Associate Editor of IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Computational Social Systems, Journal of Alzheimer's Disease, and PLoS ONE.

Zhidong Cao received the Ph.D. degree in geographic information science from the Institute of Geographic Sciences and Nature Resources Research, Chinese Academy of Sciences, Beijing, China, in 2008. He is currently an Associate Professor at the Institute of Automation, Chinese Academy of Sciences. His current research interests include social computing, public health, emergency management, and spatial analysis.

Helai Huang Dr. Helai Huang is a professor of transportation engineering in the School of Traffic and Transportation Engineering,Central South University (CSU). He serves as associate dean of the School, the founding director of CSU Urban Transport Research Centre and Key Laboratory of Smart Transportation of Hunan Province. Dr. Huang holds B.E and M.S. degrees from Tianjin University, China (1996-2003), a PhD degree in transportation engineering from National University of Singapore (2003-2007), and worked as a post-doc at University of Central Florida (2008-2010).

His research interests include traffic safety, transportation planning and ITS. Dr. Huang is a Chief Editor of Accident Analysis and Prevention (Elsevier), an editorial board member of Analytic Method in Accident Research (Elsevier) and Journal of Geography and Regional Planning, a committee member for TRB ABJ80 Committee (Statistical Methods) and TRB ANF10 Committee (Pedestrians), and a journal reviewer for over 20 SCI/SSCI -indexed international journals.

Daniel Zeng received the B.S. degree in economics and operations research from the University of Science and Technology of China, Hefei, China, and the M.S. and Ph.D. degrees in industrial administration from Carnegie Mellon University. He is Gentile Family Professor in the Department of Management Information Systems at the University of Arizona. He also holds a Research Fellow position at the Institute of Automation, Chinese Academy of Sciences. His research interests include intelligence and security informatics, infectious disease informatics, social computing, recommender systems, software agents, and applied operations research and game theory. He has published more than 300 peer-reviewed articles. He currently serves as the editor in chief of ACM Transactions on MIS, and President of the IEEE Intelligent Transportation Systems Society. He is a Fellow of IEEE.