Urban Analytics and City Science

EPB: Urban Analytics and City Science
2022, Vol. 0(0) 1–17

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23998083221107019
journals.sagepub.com/home/epb
SAGE

Spatial variations of the third and fourth COVID-19 waves in Hong Kong: A comparative study using built environment and socio-demographic characteristics

Zidong Yu and Xintao Liu
Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Abstract
Since the first confirmed case was reported in January 2020, Hong Kong has experienced multiple waves of COVID-19 outbreaks. Recent literature has explored the spatial patterns of disease incidence and their relationships with the built environment and demographic characteristics. Nonetheless, few studies aim at the comparative patterns of different epidemic waves occurring in the same spatial context. This study analyses spatial patterns of the third and fourth COVID-19 epidemic waves and then evaluates the spatial relationship between case incidence and built environment and socio-demographic characteristics. By collecting local-related cases, this study incorporates a two-fold analytical strategy: (1) Using rank-size distribution and log-odd ratio to depict the spatial pattern of COVID-19 incidence rates; (2) through global and local regression models, investigating incidence's associations with the urban built environment and socio-demographic characteristics. The results reveal that the two different epidemic waves have far distinct spatial tendencies to their infection risk factors, reflecting location-specific associations with the built environments and socio-demographics. Collectively, we discover that the third and fourth COVID-19 waves are likely associated with residential context and urban activities, respectively. Practical implications are discussed that would be of interest to policymakers and health professionals.

Keywords
COVID-19 epidemic waves, spatial analysis, built environment, socio-demographics, location-specific association

Corresponding author:
Xintao Liu, Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, 181 Chatham Road South, Kowloon, Hong Kong.
Email: xintao.liu@polyu.edu.hk

Introduction

As first discovered in China in December 2019, Coronavirus disease 2019 (COVID-19) has been recognized as an emergent contagious disease caused by the severe acute respiratory syndrome coronavirus infection. Subsequently, the World Health Organization (WHO) documented the spread of COVID-19 as a global pandemic in March 2020 due to the quickly and widely spreading to other countries (WHO, 2020). Statistics observed by health officials have reported that a significant proportion of cases are confirmed in urban areas with dense populations and affluent human activities (Liu, 2020; Paul et al., 2020). By bridging the previous outbreaks in cities, it is thus suggested that the geographical context at the urban scale can play critical roles in the course of the COVID-19 pandemic.

Existing studies have substantially contributed to the spatial patterns of infectious diseases and intimated that it is recognized as a diffusion phenomenon within a specific region (Loon, 2005). This can be attributed to the designated nature of virus transmission via the physical human-to-human interactions occurring in a shared space (Meade, 2014). Such transmissions in urban areas are considered as a more complex spatial process with multiple determinants, including a series of characteristics of built environments and socioeconomic demographics (Dietz et al., 2020; Mollalo et al., 2020). As patterns of virus spread can be substantially different across urban areas of a city, knowledge from the geospatial perspective in cities should be a necessary and powerful tool for understanding COVID-19 proliferation.

Regarding COVID-19 outbreaks reported in urban areas, the transmission process of the virus could be enabled and embedded by the spatial medium and associated mainly with built environments and socio-demographic characteristics (Drefahl et al., 2020; Huang et al., 2020; Mogi and Spijker, 2021). As for the former type of characteristics, the densely built environments in cities have driven the virus transmission towards a more sophisticated and intensive process. Explicitly, the built environment in urban settings is referred to as spatial carriers. Physical infrastructures provide the potential for human contacts and thus the precondition for the transmission of viruses. Various communities characterized by built environments may produce different impacts in strengthening the severity of local spreads. Some experiments have suggested that built environmental features (e.g., transport accessibility, land use configuration) can significantly correlate to the spatial patterns of COVID-19 contagion in cities of Asia and Europe (Huang et al., 2020; Yip et al., 2021). Meanwhile, the socio-demographic status of city residents also plays a considerable role in framing uneven geographical patterns of the COVID-19 outbreaks. Relevant literature has emphasized that communities with marginalized socio-demographic conditions may have a higher risk of being exposed to viruses. Specifically, disadvantaged socio-demographic status, including ethnic minorities, poor education, and low income, can induce residents to contract the disease more easily (Borjas, 2020; Emeruwa et al., 2020; Guo et al., 2022). Besides, the marginalized socio-demographic conditions likely lead to poor housing conditions with insufficient indoor ventilation (Ahmad et al., 2020). Although discussions of the environmental and socio-spatial issues are increasingly prevalent in research fields, a particular focus related to COVID-19 ignored mostly by current studies is the comparative patterns of different epidemic waves occurring in the same spatial context.

As aforesaid, this omission can be problematic and challenging because many countries and regions have witnessed two-wave or even multi-wave COVID-19 outbreaks. It is requisite for epidemiological professionals to confirm whether spatial shifts can be documented over different waves of outbreaks within the same spatial context. If so, the following focal points are whether and what built environment and socio-demographic characteristics can be correlated to different outbreak waves of COVID-19 in a city. Although several experiments have been proposed on the relationships between the features mentioned above and COVID-19 transmission, studies are still seldom seen from a two-wave perspective. Using Hong Kong as a study area, we, therefore, seek to bridge this research gap by asking the following questions: (1) How to identify the spatial distributions of the third and fourth epidemic waves in Hong Kong? (2) Whether spatial heterogeneities of these two-wave epidemic patterns vary over time? (3) Given the various built environment and socio-demographic characteristics, how are these variables comparatively related to the different COVID-19 spatial patterns? (i.e., what selected features will associate the exposure risk in different COVID-19 outbreaks?)

This study investigates the spatial patterns of the two-wave COVID-19 epidemic and the associated built environment and socio-demographic characteristics. First, we depict the spatial distribution of patients infected with COVID-19 during the third and fourth epidemic waves. Two metrics, including rank-size distribution and log-odds ratio, are implemented to quantify and compare spatial heterogeneity. Then, ordinary least squares regression (OLS) and geographically weighted regression (GWR) models are applied to assess the relationships between the two-wave patterns and the characteristics of the associated built environment and socio-demographics.

Literature review

The spatial associations between the incidence of epidemic and local factors have long raised the concerns of scholars from multi-discipline research communities. Substantial evidence from previous studies has shown that disease incidence of an epidemic can be associated with different locational contexts, with much of this variation explained by characteristics like population density, age level, urban structure, race composition, and others (Drefahl et al., 2020; Kwok et al., 2020). Literature related to spatial variation of COVID-10 cases has also highlighted that significant differences can be observed across geographical units, such as city, county, or neighborhood (Kuebart and Stabler, 2020).

The spread of infectious diseases is not only the consequence of biological features of the virus but also a manifestation embedded by intensive human contacts within physical space. Spatial analyses have been performed to learn underlying connections between disease incidence and local variables of built environment characteristics. Some have confirmed that the environment plays a considerable role in disease dynamics and the health conditions of individuals (Emeruwa et al., 2020; Spencer et al., 2020). Pinter-Wollman et al. (2018) highlighted that the built environment can exhibit a strong relationship with the containment of infectious diseases, for instance, the built structures may shape the complexity of infectious diseases. These built structures can be characterized as the urban geometry attributes such as building geometry, sky view factor (SVF), road networks, and the distribution of green spaces. Similar findings are also reported by Kwok et al. (2020) that applied urban geometry characteristics to explain the distribution of local COVID-19 cases confirmed in Hong Kong, suggesting that building geometry and road network settings could be factors to increase spreading risk. Compared to rural regions, research by Lee et al. (2020) summarized that urban settings have some common environmental characteristics, which are critical factors in facilitating the virus transmission process. In the US, for example, Paul et al. (2020) emphasized that the mean prevalence of COVID-19 in urban counties (107.6 per 100,000) is significantly higher than in rural counties (43.6 per 100,000) within the same period. Mogi and Spijker (2021) further reported that population density is strongly and positively associated with COVID-19 in European countries. Therefore, it is suggested that one built environment factor in urban areas that is highly related to COVID-19 spread is density. In this sense, some evidence has been proposed to demonstrate that the high density of commercial venues with intense face-to-face contacts can lead to the harmful behavior of contracting infectious viruses (Liu, 2020; Yip et al., 2021). Collectively, it can be seen that the COVID-19 transmission process in cities to some extent correlates to the specific settings of urban built environments.

In addition to built environments, another stream of literature has reported that the socio-demographic status of individuals can also contribute to the incidence of epidemics. The range of socio-demographic characteristics is broad and covers comprehensive facets of the COVID-19 pandemic. Almagro and Orane-Hutchinson (2020) provided helpful insight to measure the importance of socioeconomic factors such as population density, commuting patterns, and occupations in explaining the spatial disparities of COVID-19 cases. A crucial implication drawn from their investigation is that patients' occupations are testified as a critical factor in explaining this epidemic's early transmission across NYC neighborhoods. Borjas (2020) claimed that poor or immigrant neighborhoods with disadvantaged socio-demographic status have lower opportunities for virus tests and higher risks to be confirmed as positive cases. One interesting point is that these marginalized neighborhoods contain larger household or predominantly black populations, illustrating a medical inequality caused by race composition regarding COVID-19 risk. To further explain this phenomenon, some have figured out that people residing in communities with poor socioeconomic developments may have worse working environments, lower levels of health awareness, and more frequent usage of public transportation (Chen, 2020; Sá, 2020). For instance, blue-collar workers at worsening socio-demographic status are faced with scarce health protection in precarious environments like slaughterhouses or construction sites. By in turn linking to the built environments, Lee et al. (2020) suggested that such physiological and medical inequalities in marginalized neighborhoods are more likely to be exacerbated by denser housing space, insufficient supplies of freshwater, and poor sanitation facilities. In this context, populations with such socio-demographic characteristics may have a higher exposure risk to the COVID-19 outbreak.

The ongoing pandemic of COVID-19 has progressively caused severe threats to most countries, in which many of them have experienced multiple waves of reported cases to date. With that being said, we still lack specific knowledge of the spatial variation patterns of a two-wave COVID-19 epidemic. That is, little effort has been devoted to quantifying the collective differences of several statistical peaks of epidemics, with respect to the surrounding built environment and local socio-demographics. Saito et al. (2020) presented a brief comparison of the severity and patience characteristics of the first and second waves in Japan, while Kan et al. (2021) introduced a space-time scan statistic technique to detect the high-risk areas of COVID-19 cases reported in the first three outbreak waves in Hong Kong.

Study area and dataset

To investigate the patterns of two-wave COVID-19 outbreaks, Hong Kong, a Specialized Administrative Region (SAR) of China, is designated as the study area. It covers a total area of 1100 km^2 with a total population of more than 7 million people. We consider Hong Kong as a reasonable case study for the following reasons. First, since the initial COVID-19 case was reported on January 23rd, 2020, Hong Kong has experienced four waves of epidemic outbreaks with over 11,100 confirmed cases by March 8th, 2021, which can be primarily classified into imported and local-related cases. Second, Hong Kong is one of the most densely populated cities in the world. Most residents live in a built-up area with approximately 264 km^2, accounting for 24% of the total area. A small proportion of the area is seen as the urbanized regions for human activities, in which a high urban density is thus anticipated and provides a representative urban environment for inspecting the transmission of the COVID-19 virus. The built environment and socio-demographic features are disparate in this city, ranging from well-developed Central Business Districts to high-density residential neighborhoods.

As shown in Figure 1(a) and suggested by supplementary table S1, our research scope is the third and fourth wave epidemics, which are regarded as two separate phases: Phase I between 6th July and 28th August 2020, while Phase II between 8th November 2020 and 20th February 2021.

The COVID-19 data used in this research is retrieved from the Department of Health in Hong Kong, which is freely accessed at https://data.gov.hk/en-data/dataset/hk-dh-chpsebcddr-novel-infectious-agent. This dataset depicts the details of confirmed cases, including reported and onset dates, residential location, visit place, case classification, and some individual information (Supplementary figure S1). The COVID-19 cases reported in Hong Kong can be classified into two categories in accordance with the origin of infection, namely, imported-related cases that are patients who were infected overseas, and local-related cases, which include cases with clearly local sources or untraceable sources (i.e., local cases), and cases involving infection by local cases (i.e., Epidemiologically to local cases). For spatial analysis, all residential locations are considered and geocoded for their coordinates that can be aggregated into DCCAs, while Figure 1(b) exhibits the spatial distribution of the confirmed cases.

A wide range of built environment and socio-demographic characteristics has been demonstrated in associating with the COVID-19 transmission process in cities (Huang et al., 2020; Yip et al., 2021). Accordingly, this paper hypothesizes that the urban environments and socio-demographics carry relatively evident relationships with virus transmission. As suggested by Ewing and Cervero (2010), we conceptualize built environment indicators into 5Ds (i.e., density, diversity, design, destination accessibility, and distance to transit). Specifically, density reflects the activity density of a place that has often been calculated by population and building numbers per spatial unit. Diversity indicates the intensity of miscellaneous urban functions and services, while design assesses the urban structure, such as road networks and pedestrian-oriented infrastructure (Zhang et al., 2021). Destination accessibility evaluates the ease of access to trip attractions, and distance to transit is often considered as the distance from the residences to the nearest public transit hubs. Besides, socio-demographics are also recommended as important factors to explain the incidence of COVID-19 and thus implemented in this spatial analysis. Retrieved from the 2016 population by-census, four aggregated metrics are implemented involving education level, population age level, mean income, jobs-housing ratio, and race composition. Supplementary table S2 summarizes the built environment and socio-demographic characteristics applied in our case.

Methodology

Rank-size distribution and log-odds ratio

Rank-size distribution has been widely applied as one of the most useful tools for assessing the heterogeneity of human concentrated patterns founded on their magnitudes (Reed, 2002). Mathematically, rank-size order manifests the frequency distribution of collective events. Here, we implement rank-size distribution to report the spatial heterogeneity of case concentration and investigate how the spatial distribution patterns vary between two phases. The COVID-19 intensity is defined using the incidence rate, which is the number of reported cases per 1000 people in each of DCCAs during a particular time period:

$$In_j(d_i) = F_j(d_i) \big/ Pop(d_i) \tag{1}$$

Here, during a phase j, $In_j(d_i)$ represents the incidence rate of COVID-19 cases in a DCCA (d_i),$F_j(d_i)$, is the number of confirmed cases located in a DCCA d_i, and $Pop(d_i)$ denotes the population $(*1000)$ in the same DCCA d_i. Following this logic, given a phase, a rank-size distribution can be generated to reflect the spatial heterogeneity and concentration of COVID-19 incidence rate. It should be noted that incidence rate only includes new cases during a certain period while the prevalence rate includes new and pre-existing cases.

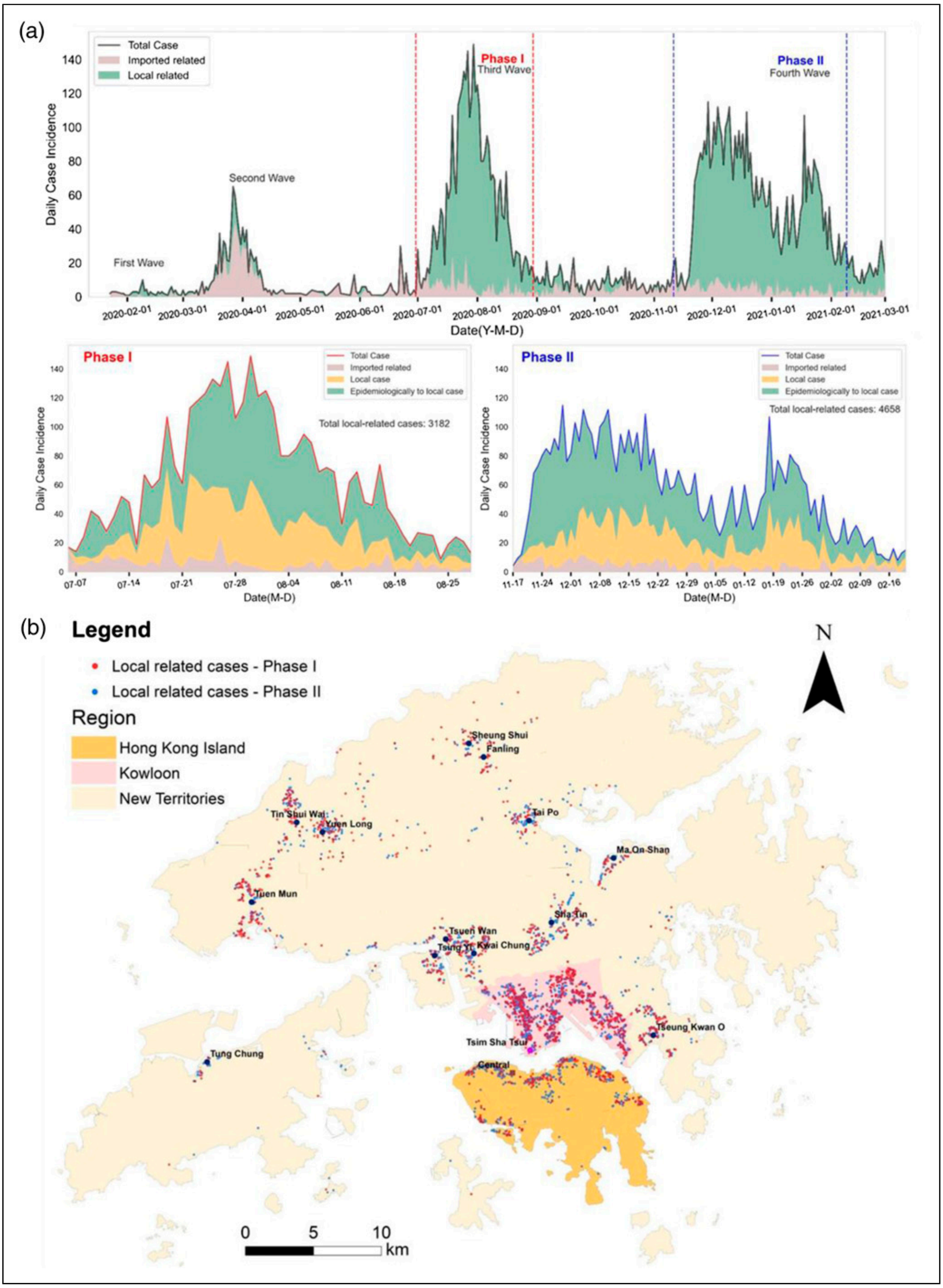

Figure 1. (a) Temporal distribution of COVID-19 cases reported in Hong Kong. (b) The spatial distribution of the confirmed COVID-19 cases in two separate phases at the level of District Council Constituency Area (DCCA) with established urban centers and new towns in New Territories.

Although using rank-size distribution can help us capture a holistic image of how COVID-19 incidence of different phases is distributed across space, it does not link the two phases on the same geographic scale. As such, we consider another evaluation metric, log-odds ratio, to depict the relative balance of phase I and phase II incidences. Given that the total incidence of phase I and phase II are different, a normalization process based on the total number of confirmed cases of different phase $(i.e., T_j \, and \, T_k)$ is requisite to support the comparison. In a DCCA d_i , the log-odds ratio of two phase j and k, is represented as:

$$R(d_i) = \log_{10}\left(\frac{In_j(d_i) \cdot T_j}{In_k(d_i) \cdot T_k}\right) \tag{2}$$

The value of $R(d_i)$ can be categorized into three interpretations. A value of $R(d_i)$ equals to zero indicates that the relative incidence of COVID-19 is identical in area d_i, while a value larger than zero denotes that the relative incidence of phase j is higher than that of phase k, and vice versa.

Global and local regression analysis

We then incorporate both global and local regression models to explore the association of the normalized size of COVID-19 incidence with the built environment and socio-demographic characteristics. It is worth noting that the variance inflation factor (VIF) is used to assess the correlations among all independent variables in addressing multicollinearity issues.

The aim of OLS regression is to identify and measure relationships examine the relationship between the incidence of COVID-19 and built environments and socio-demographic factors. The formula of OLS is illustrated in equation S1. This model can provide a holistic image of the examined relationship and serve as a baseline model in this study.

As the independent variables are probably related to the local context, the spatial non-stationarity of study variables can be ignored by using a global regression model (Crespo and Grêt-Regamey, 2013). Geographically weighted regression (GWR) is hence introduced to tackle the issues of spatial non-stationarity, which can be regarded as a varying association between independent and dependent variables (Wang et al., 2018). GWR is proposed as a local regression model, allowing the variable coefficients to be varied across space and mitigating the statistical bias caused by spatial non-stationary. The equation of GWR is shown follows

$$In_j(d_i) = \beta_{0i} + \sum_k \beta_{ki} x_{ki} \tag{3}$$

where In_j is the COVID-19 incidence rate in phase j, β_{0i} represents the local regression intercept, and x_{ki} and β_{ki} denote the independent variables and the local estimated regression coefficients of DCCA i, separately. It should be mentioned that β_{ki} is a spatial function and its values are associated with the results of applying weights on geographic coordinates of the observation DCCA i. In this case, we select the Gaussian weighting function to define the weighting structure since it can provide a continuous function for distance and thus better calibrate the model. Akaike Information Criterion (AIC) is tested to define the optimal kernel bandwidth.

Analytical results

Spatial patterns of the COVID-19 incidence

In reply to the RQ1 related to the spatial geographic distribution of COVID-19 incidence, we provide rank-size distribution and log-odds ratios for exploring the spatial patterns in this subsection. As manifested in Figure 2(a), the epidemic incidence from both phases decays dramatically for the first few DCCAs, yielding that these areas contain a substantial amount of confirmed cases with intensive epidemic situations. From a statistical perspective, this phenomenon can be described as the king effect, where the percentage of the top-ranked observations is much larger than the rest ones. It depicts a high degree of spatial heterogeneity of COVID-19 cases in respect to both phases at the DCCA level. As the rank increases, it is clear to observe that the disease incidence of the remaining areas tends to decay more smoothly compared to the first few areas. In order to quantify the decay patterns, we furthermore fit the curves using the power-law function shown as $y = a \cdot x^b$, in which the parameter b reflects the rate of decay. Fitted curves for phase I and phase II are $y = 5.43 \cdot x^{-0.52}$ and $y = 9.12 \cdot x^{-0.57}$, while the corresponding R squared values are 0.95 and 0.96, respectively. It should be stressed that the decay rates of rank-size distribution are similar to each other, implying that decaying speed is comparable between two different phases.

Figure 2(b) exhibits the spatial distribution of disease incidence with the top-ranked areas for phase I and Phase II. From a visual perspective, the two phases show relatively disparate patterns. For example, in phase I, the top 2 ranked DCCAs are clustered in Wong Tai Sin, a high-density

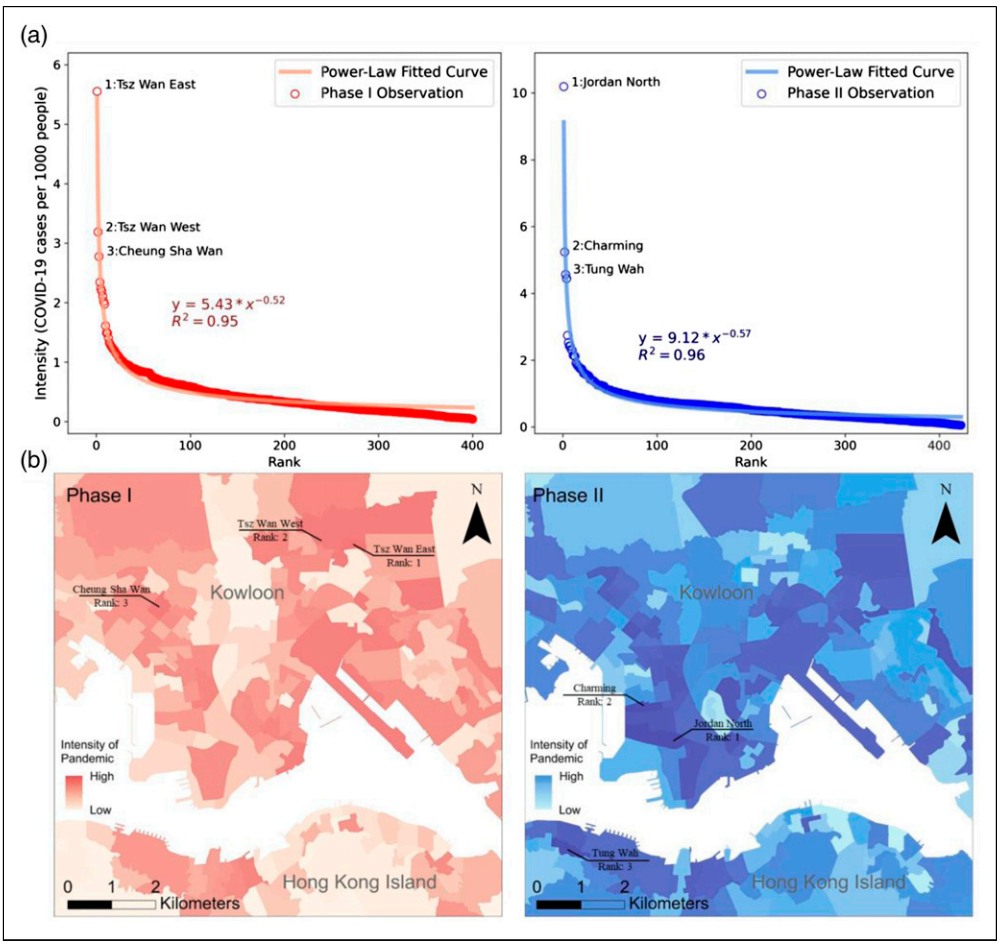

Figure 2. (a) Rank-size distribution of the COVID-19 incidence of the two phases. (b) Spatial distribution of the COVID-19 incidence of the two phases with the top 3 ranked DCCAs.

residential district in northeast Kowloon, with more than 85% of the district's residents living in public housing estates. Conversely, during phase II, DCCAs with a high level of COVID-19 incidence tend to situate at Jordan, a mixed-function neighborhood located in the Kowloon Peninsula with commercial, residential, entertainment services for locals and tourists.

Despite that the statistical and spatial distributions have provided concentration patterns, a more direct comparison linking two incidence layers in the same geographic regions is requisite. A log-odds ratio between phase I and phase II is computed here, showing the spatial variation of epidemic patterns shift between two distinct phases. Supplementary figure S2 displays the histogram of log-odds ratios and its spatial distribution, presenting an overview of the relative balance as regards two-wave COVID-19 incidence over space. DCCAs with opposite ratios produce similar distributions, from where 49.2% and 50.8% of DCCAs are observed with positive and negative ratios, respectively. This result highlights that the COVID-19 concentration of those two phases is mostly equivalent in Hong Kong, following the findings of the overall rank-size distribution. We additionally derive some interesting patterns from the geographic distribution. During phase I, a higher level of COVID-19 incidence rate is reported in the peripheral areas of the downtown core and new towns in New Territories, such as Wong Tai Sin, Tuen Mun, and Fanling. In contrast, the central part of Hong Kong situated in the northwestern part of Hong Kong Island and the south shore of Kowloon Peninsula exhibit more severe situations during phase II. Plus, a similar pattern is observed in Causeway Bay in eastern Hong Kong Island, which is seen as an energetic trade and commercial hub in Hong Kong. These findings thus can establish that downtown areas with vibrant daily activities may be correlated to the risk of virus transmission in phase II.

The relationships with built environment and socio-demographic characteristics

OLS regression results. The OLS regression model is used to examine the global relationships between the COVID-19 disease incidence and built environment and socio-demographic characteristics. Supplementary table S3 reports the regression results of the OLS regression model, including estimated coefficient, standard error, VIF, adjusted R-squared, and AIC.

Variables from phase I outbreak, including POI diversity, distance to Central, population age level, mean income, jobs-housing ratio, and race composition of Filipino are significantly related to the COVID-19 incidence. Among these variables, mean income indicates the strongest negative relationship with a coefficient of −0.24, implying that areas with lower income levels may produce higher exposure risk to the COVID-19 virus. Research by Chen (2020) has suggested that lower household income can subject socially marginalized and precarious communities to heightened exposure to infectious disease. Housing characteristics can be regarded as one of the essential reasons that densely populated neighborhoods of public housing and cramped conditions of living environment enhance the spread process of COVID-19. The coefficient of jobs-housing ratio also indicates that for the phase I outbreak, residential districts with fewer local working opportunities tend to be hotspots of virus spread. Compared to phase I, the regression results from phase II hint at a relatively dissimilar association to the selected local variables. Explicitly, POI density exhibits a positive correlation with disease incidence, suggesting that areas with intensive human activities probably contain higher COVID-19 risk. This research finding is also reported by previous literature that hotspots and sub-centers in downtown areas are activity hubs, including populated human interactions and daily activities. In an epidemic context, such hotspots can be potential virus spreaders since visits and contacts can increase the risk of transmission of COVI-19 disease (Chang et al., 2020). Additionally, the variable of distance to Central shows a substantially negative association with the disease incidence. Although the variables mentioned above are reported with significant associations, the values of adjusted R-squared suggest a limited capacity in explaining the variation of the incidence rate in DCCAs.

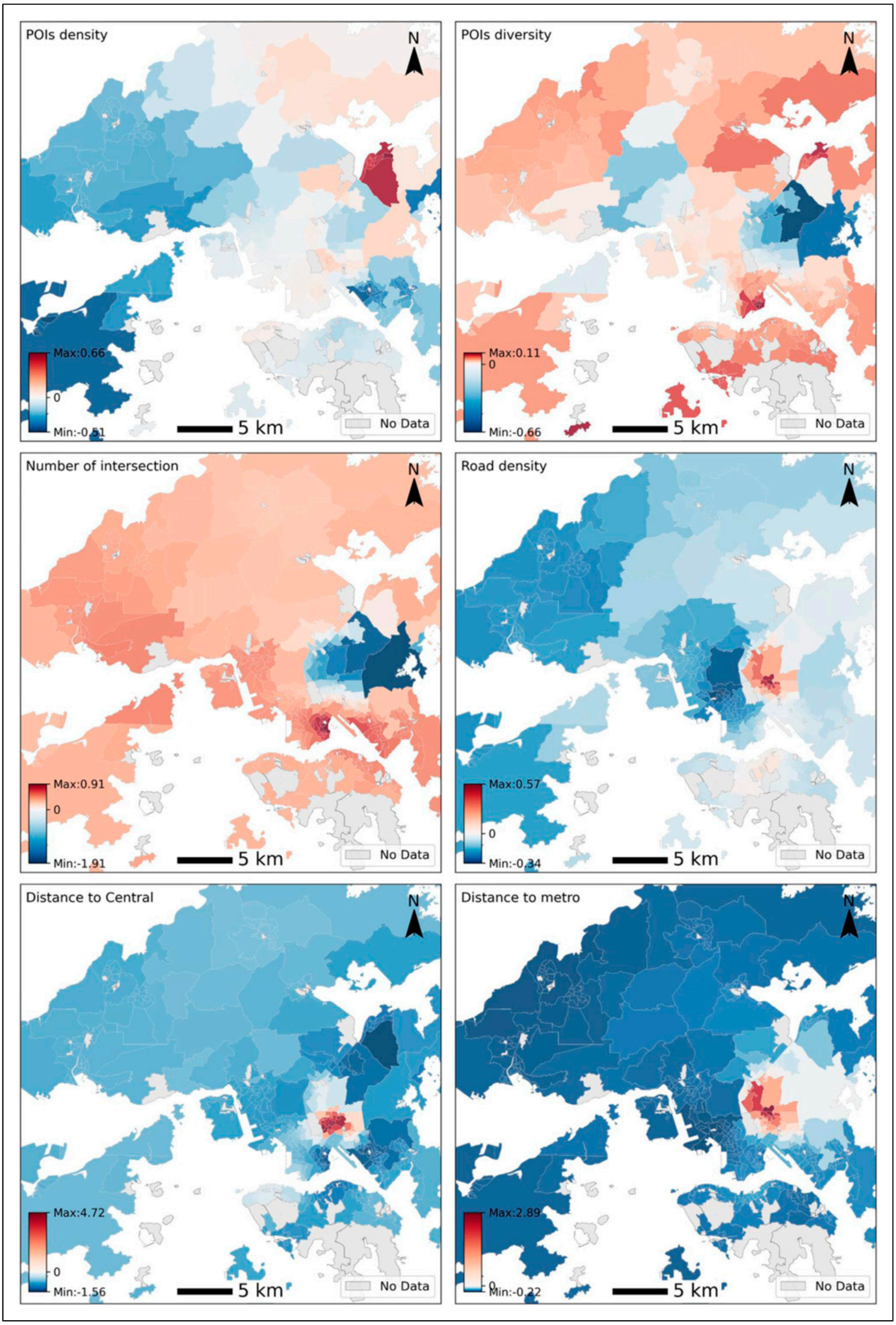

Figure 3. Spatial distribution of associated regression coefficients (built environment characteristics) for phase I outbreak.

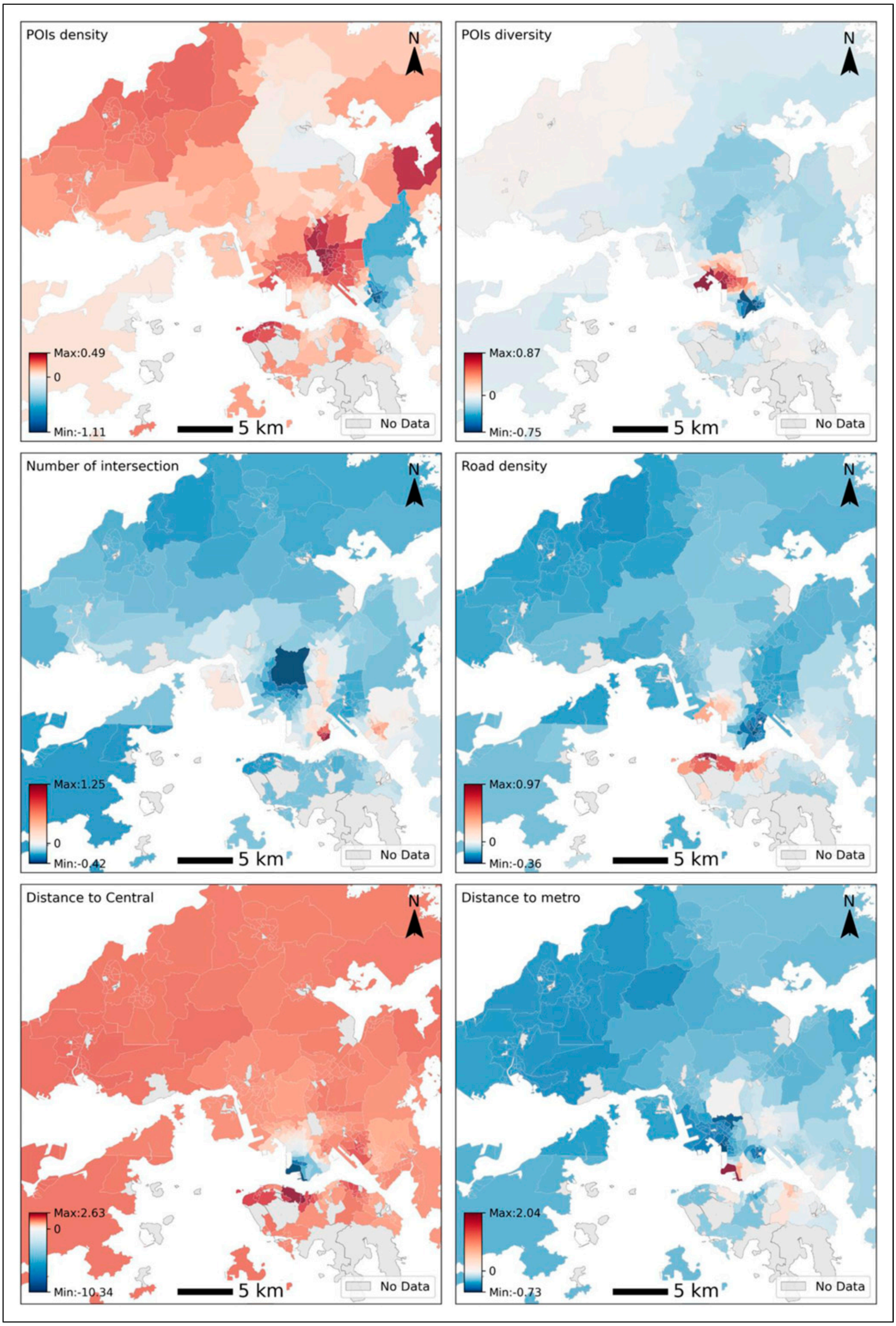

Figure 4. Spatial distribution of associated regression coefficients (built environment characteristics) for phase II outbreak.

GWR results. The results from global OLS regression may not consider some variation related to the relationships across geographic space (i.e., spatial non-stationary). We therefore present the results generated from GWR models, which can be used to explore the spatial non-stationary We present the results generated from GWR models, which focus on the spatial distribution of the significantly associated local characteristics. As reported by output results shown in supplementary table S4, the R-squared values of phase I and phase II GWR models are 0.47 and 0.55, respectively, which are considerably higher than that of the OLS models. Supplementary figure S3 exhibits the spatial distributions of local R-squared of the GWR models. It can be evident that the GWR models of phases I and II can produce relatively higher explanatory capacity in the eastern parts of the study area and northern Hong Kong Island, separately. The AIC values of GWR models are lower than those of OLS models, which are 1058.21 and 1052.65 in GWR compared to the 1060.31 and 1161.65 in OLS models for phase I and Phase II.

As reported in Figure 3, positive coefficients of road density, distance to Central and metro stations are mostly located in Wong Tai Sin, where POIs diversity and the number of intersections display opposite spatial patterns. One anticipated finding is that the variable of distance to Central exhibits an obvious positive relationship with the case incidence, suggesting that the incidence of COVID-19 tends to be observed in peripheral areas during phase I. As for the socio-demographics shown in supplementary figure S4, variables except for the percentage of Filipino indicate a robust negative relationship in northeastern Kowloon, where the Wong Tai Sin district is located. Compared to other non-residential districts, the findings in northeastern Kowloon are meaningful. Since mean income and age level can represent important profiles of residential neighborhoods, the result may imply that these two socio-demographic characteristics are critical infection risk factors, which are highly related to housing environments. The research findings demonstrated here are also reported by Kwok et al. (2020) that deficient urban design and geometry in Hong Kong are significantly related to the transmission of viruses. In other words, people living in old residential estates or public housing can gain a higher risk of being infected with COVID-19 disease. Moving to the phase II outbreak, we discover some interesting patterns in southern Kowloon and northern Hong Kong Island. Such areas are regarded as the activity centers of Hong Kong. To be specific, Figure 4 manifests that the associated built environment characteristics have a formidable relationship on the Kowloon Peninsula. For instance, in most areas, POIs density presents positive relationships with local incidence rates. It is therefore suggested that in these areas, COVDI-19 risk may be positively connected to residents' daily activities. Correspondingly, results from the former subsections also highlight places as the spreading hotspots during phase II outbreak. Supplementary figure S5 reports the spatial coefficients distribution of socio-demographic characteristics. Among these factors, one interesting finding is that the incidence rate in the eastern Kowloon Peninsula exhibits an evident pattern with strong positive correlations with the jobs-housing ratio in comparison with remaining areas. This result may be explained by the fact that neighborhoods in this area (e.g., Kowloon City) mostly provide mixed commercial and residential functions, including public housing estates and multiple day-to-day commercial services.

Discussion and conclusion

The spatial variation of epidemic incidence and its relationship to local factors have been long studied in research communities. Recently, in response to the COVID-19 epidemic outbreak, literature has provided substantial evidence to the knowledge of virus transmission, but very few aims at emphasizing spatial variations of epidemic waves occurring in the same geographical context. As this study shows, we compare spatial patterns of the third and fourth COVID-19 waves in Hong Kong and assess the relationships between COVID-19 incidence rate and built environment and socio-demographic characteristics.

By leveraging rank-size distribution, we quantify the spatial patterns of COVID-19 incidence at the DCCA level. It is apparent from this distribution that disease incidence has a notable decay effect on both the third and fourth waves. This means that a considerable number of COVID-19 cases have a propensity to concentrate in a few areas of the city. The concentrated patterns may be explained by the fact that the epidemic outbreak is implemented through absolute space featured by concrete spatiality, hinting that the COVID-19 transmission is enabled by people sharing the same physical environment of local communities (Moore et al., 2020). This specific feature of infectious viruses makes disease incidence more concentrated within certain areas. Furthermore, the spatial distribution of disease incidence with the top-ranked areas is produced. This key finding highlights that the high-risk hotspots of virus transmission within one epidemiological period can be spatially clustered at the aggregate level.

As the rank-size distribution cannot directly compare two epidemic outbreaks in the same geographic space, we take further steps by adopting a log-odds ratio to reflect the relative balance of COVID-19 incidence. DCCAs with opposite values tend to share a similar proportion of the entire study area, showing that the comparative distribution of both epidemics is even and comparable. An overrepresentation of relative incidence in Wong Tai Sin district suggests that this residential district is the hardest-hit area during the phase I outbreak. In contrast, the mismatch of spatial patterns observed from the downtown areas intimates that a more critical situation is shown in urban centers with vibrant human activities during the phase II outbreak. Such areas have long been recognized as the commercial and political center of Hong Kong (Yu and Liu, 2021). The spatial shifts of spreading hotspots deliver critical evidence to illustrate that COVID-19 outbreaks may be highly correlated to location-specific characteristics (Franch-Pardo et al., 2020; Kan et al., 2021). In addition, the log-odds ratio provides an intuitive metric to visualize the spatial heterogeneities of incidence rate between two epidemic waves.

This study then reports some remarkable results by investigating how built environmental and socio-demographic local factors related to the disease incidence at an aggregate level. OLS regression is first implemented to inspect the independent variables that may be related to explaining the variation of COVID-19 incidence. Results reveal that correlated characteristics in different waves elucidate dissimilar tendencies. More specifically, the relationships identified from the third wave exhibit evident residential-oriented features, of which mean income, JHR, and race composition of white and Filipino population are highly associated with residential neighborhoods. In this context, local marginalized communities with lower economic status face a higher risk of COVID-19 outbreak that can be attributed to the densely populated public housing and cramped conditions of the living environment (Ahmad et al., 2020). Apart from the third wave, the fourth wave of epidemic outbreaks presents disparities as regards the infection risk factors. POIs density proposes a positive relationship with epidemic incidence, highlighting that places with urban vibrancies presumably produce a higher contact risk of COVID-19 virus to local neighborhoods. Such places can be potential super-spreading venues as visited by a greater volume of population flows, which can raise the chance of contracting infectious diseases (Chang et al., 2020; Tenforde et al., 2021). The GWR model is then applied to examine the spatial non-stationarity by understanding how the relationships between disease incidence and local characteristics vary by geographic space. From the spatial distribution of estimated coefficients, we can easily notice that hotspot areas for each epidemic wave display a significant effect in predominating the coefficient magnitude, in which these areas are in agreement with those obtained by previous results. This empirical thesis has provided a deeper insight into the local characteristics that are closely related to infectious disease incidence. Taken together, this study presents a global regression analysis in conjunction with a local perspective to enhance the understanding of how selected variables are related to the COVID-19 incidence in Hong Kong. The global approach seems to be more suitable in offering holistic evidence for revealing universal relationships and thus discussing regional-wide policies, while the results from local analysis can report more nuanced and complementary information (Crespo and Grêt-Regamey, 2013).

Some sources of weakness should be acknowledged. First, residential locations of confirmed cases are used to indicate the distribution of COVID-19 incidence. We assume that COVID-19 patients contract the COVID-19 virus within or near their residential neighborhoods. However, human commuting behaviors that visit public spaces like offices or restaurants in other districts can increase the risk of contracting COVID-19, which is not included in this study. Another major weakness is that many non-pharmacological interventions (NPIs) such as community interventions and border control policies have been demonstrated as effective measures preventing the spread of Covid-19 but are not comparatively considered in this research. Previous studies have largely suggested that NPIs can be adopted as mitigation strategies apart from medical treatments in Hong Kong (Chan et al., 2021; Yu et al., 2021). It raises new questions about such qualitative NPIs implemented by the government and should be answered in the future by collecting targeted datasets. In this respect, one plausible strategy for subsequent research is to conduct a questionnaire regarding the COVID-19 NPIs' impact on the public. Besides, we only explore the spatial aspects of COVID-19 outbreaks by analyzing the distribution patterns of observed cases and their correlations to local factors. Individual-level information such as the patient's age, gender, death, and discharge rates should be considered, which are important indicators helping us to evaluate the epidemic situation.

Practical implications for epidemic control

The findings of this study can produce some practical and policy implications that would be of interest to relevant policymakers and health professionals. Referring to the third wave (phase I), the research has presented evidence that residential districts are the hardest-hit areas, of which marginalized groups living in poor housing conditions can be closely associated with a higher risk of COVID-19 due to the deficient living environments. To prevent the spread or mitigate the severity of regional epidemics, targeted measures are therefore suggested to these residence groups. For example, community-based virus testing sites can be temporarily set up in residential neighborhoods with reported confirmed cases for providing no-cost COVID-19 testing services. Such measures may help health officials to detect silent infections in high-risk populations, thus cutting off the infectious chains. Regular cleaning and disinfection are also encouraged to maintain the hygiene of public infrastructures of communities, particularly given the old and dense properties with poor ventilation systems. Indeed, mid- and long-range airborne transmissions by ventilation issues have been documented recently as a plausible cause in triggering COVID-19 outbreaks in high-density public residential buildings in Hong Kong (Li et al., 2021; Tung et al., 2021).

As for the fourth wave (Phase II) epidemic outbreak, this study supplements some hints to the recent discussions of social gathering activities. The fourth outbreak displays a tendency towards downtown areas where vibrant social activities densely occur. Notably, gathering events in holidays such as Halloween, Christmas Day, and New Year's Eve before or during the epidemic period may increase the chance of spreading and getting COVID-19 due to potential virus exposure in populated circumstances. The evidence observed by Kuebart and Stabler (2020) has confirmed that cultural festivities and street parades during carnival events in downtown areas can facilitate the spreading process of the disease. Such activities are mostly characterized by close personal contact, food, and alcoholic consumption. It is thus recommended that restrictions on travel and social behaviors are vital, to curb non-essential gatherings, for cutting off the silent transmission link in the community in a short time. Explicitly, tighter measures on personal behavior like requiring physical distancing in face-to-face interaction or prohibiting dine-in activities could be imposed on some particular categories of places such as restaurants and bars.

Combined together, effective and targeted implementations from our implications can be to some extent used as constructive suggestions for government and the public. In this regard, practical implications emerged and discussed in this study can help to respond to the emerging transmission chains found in different neighborhoods at the beginning of the next outbreak.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Xintao Liu https://orcid.org/0000-0002-7323-9878

Supplemental Material

Supplemental material for this article is available online.

References

Ahmad K, Erqou S, Shah N, et al. (2020) Association of poor housing conditions with COVID-19 incidence and mortality across US counties. *Plos One* 15(11): e0241327.

Almagro M and Orane-Hutchinson A (2020) JUE insight: The determinants of the differential exposure to COVID-19 in New York city and their evolution over time. *Journal of Urban Economics* 127: 103293.

Borjas GJ (2020) *Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods (No. w26952)*. Cambridge, MA: National Bureau of Economic Research.

Chan HY, Chen A, Ma W, et al. (2021) COVID-19, community response, public policy, and travel patterns: a tale of Hong Kong. *Transport Policy* 106: 173–184.

Chang SY, Pierson E, Koh PW, et al. (2020) Mobility network modeling explains higher SARS-CoV-2 infection rates among disadvantaged groups and informs reopening strategies. medRxiv.

Chen AL (2020) A socio-spatial research agenda on the COVID-19 pandemic. *Acta Sociologica* 63(4): 453–456.

Crespo R and Grêt-Regamey A (2013) Local hedonic house-price modelling for urban planners: advantages of using local regression techniques. *Environment and Planning B: Planning and Design* 40(4): 664–682.

Dietz L, Horve PF, Coil DA, et al. (2020) 2019 novel coronavirus (COVID-19) pandemic: built environment considerations to reduce transmission. *Msystems* 5(2): pp.e00245-20.

Drefahl S, Wallace M, Mussino E, et al. (2020) Socio-demographic risk factors of COVID-19 deaths in Sweden: a nationwide register study. *Department of Sociology*. doi: 10.17045/sthlmuni.12420347.v3.

Emeruwa UN, Ona S, Shaman JL, et al. (2020) Associations between built environment, neighborhood socioeconomic status, and SARS-CoV-2 infection among pregnant women in New York City. *Jama* 324(4): 390–392.

Ewing R and Cervero R (2010) Travel and the built environment: a meta-analysis. *Journal of the American Planning Association* 76(3): 265–294.

Franch-Pardo I, Napoletano BM, Rosete-Verges F, et al. (2020) Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment* 739: 140033.

Guo Z, Liu X and Zhao P (2022) A vector field approach to estimating environmental exposure using human activity data. *ISPRS International Journal of Geo-Information* 11(2): 135.

Huang J, Kwan MP, Kan Z, et al. (2020) Investigating the relationship between the built environment and relative risk of COVID-19 in Hong Kong. *ISPRS International Journal of Geo-Information* 9(11): 624.

Kan Z, Kwan MP, Huang J, et al. (2021) Comparing the space-time patterns of high-risk areas in different waves of COVID-19 in Hong Kong. *Transactions in GIS* 25: 2982–3001.

Kuebart A and Stabler M (2020) Infectious diseases as socio-spatial processes: the Covid-19 outbreak in Germany. *Tijdschrift voor economische en sociale geografie* 111(3): 482–496.

Kwok CYT, Wong MS, Chan KL, et al. (2020) Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19, a study in Hong Kong. *Science of The Total Environment* 764: 144455.

Lee VJ, Ho M, Kai CW, et al. (2020) Epidemic preparedness in urban settings: new challenges and opportunities. *The Lancet Infectious Diseases* 20(5): 527–529.

Li Y, Qian H, Hang J, et al. (2021) Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. *Building and Environment* 196: 107788.

Liu L (2020) Emerging study on the transmission of the novel coronavirus (COVID-19) from urban perspective: evidence from China. *Cities* 103: 102759.

Loon JV (2005) Epidemic space. *Critical Public Health* 15(1): 39–52.

Meade MS (2014) Medical geography. *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society* 1375–1381.

Mogi R and Spijker J (2021) The influence of social and economic ties to the spread of COVID-19 in Europe. *Journal of Population Research* 1–17.

Mollalo A, Vahedi B and Rivera KM (2020) GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the total environment* 728: 138884.

Moore RC, Lee A, Hancock JT, et al. (2020) Experience with social distancing early in the COVID-19 pandemic in the United States: implications for public health messaging. MedRxiv.

Paul R, Arif AA, Adeyemi O, et al. (2020) Progression of COVID-19 from urban to rural areas in the United States: a spatiotemporal analysis of prevalence rates. *The Journal of Rural Health* 36(4): 591–601.

Pinter-Wollman N, Jelić A and Wells NM (2018) The impact of the built environment on health behaviours and disease transmission in social systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373(1753): 20170245.

Reed WJ (2002) On the rank-size distribution for human settlements. *Journal of Regional Science* 42(1): 1–17.

Sá F. (2020). Socioeconomic determinants of Covid-19 infections and mortality: evidence from England and Wales. CEPR Discussion Paper No. DP14781.

Saito S, Asai Y, Matsunaga N, et al. (2020) First and second COVID-19 waves in Japan: a comparison of disease severity and characteristics: comparison of the two COVID-19 waves in Japan. *The Journal of Infection* 82(4): 84–123.

Spencer JH, Finucane ML, Fox JM, et al. (2020) Emerging infectious disease, the household built environment characteristics, and urban planning: Evidence on avian influenza in Vietnam. *Landscape and Urban Planning* 193: 103681.

Tenforde MW, Fisher KA and Patel MM (2021) Identifying COVID-19 risk through observational studies to inform control measures. *JAMA* 325: 1464.

Tung CW, Mak CM, Niu JL, et al. (2021) Enlightenment of re-entry airflow: the path of the airflow and the airborne pollutants transmission in buildings. *Building and Environment* 195: 107760.

Wang J, Lee K and Kwan MP (2018) Environmental influences on leisure-time physical inactivity in the US: An exploration of spatial non-stationarity. *ISPRS International Journal of Geo-Information* 7(4): 143.

World Health Organization (2020) WHO director-General's opening remarks at the media briefing on COVID-19-11 March 2020. Available at: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed 2 March 2021).

Yip TL, Huang Y and Liang C (2021) Built environment and the metropolitan pandemic: analysis of the COVID-19 spread in Hong Kong. *Building and Environment* 188: 107471.

Yu Z and Liu X (2021) Urban agglomeration economies and their relationships to built environment and socio-demographic characteristics in Hong Kong. *Habitat International* 117: 102417.

Yu Z, Zhu X, Liu X, et al. (2021) Reopening international borders without quarantine: contact tracing integrated policy against COVID-19. *International Journal of Environmental Research and Public Health* 18(14): 7494.

Zhang J, Liu X, Tan X, et al. (2021) Nighttime vitality and its relationship to urban diversity: an exploratory analysis in Shenzhen, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 15: 309–322.

Author biographies

Zidong Yu a PhD student in the Department of Land Surveying and Geo-Informatics in Hong Kong Polytechnic University. His research interests include urban informatics, and geospatial analysis on socioeconomics.

Xintao Liu an Assistant Professor in the Department of Land Surveying and Geo-Informatics in The Hong Kong Polytechnic University. His research interest focuses on GIScience, transportation geography, and complex network.