Science of the Total Environment 772 (2021) 145379

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Identifying the space-time patterns of COVID-19 risk and their associations with different built environment features in Hong Kong

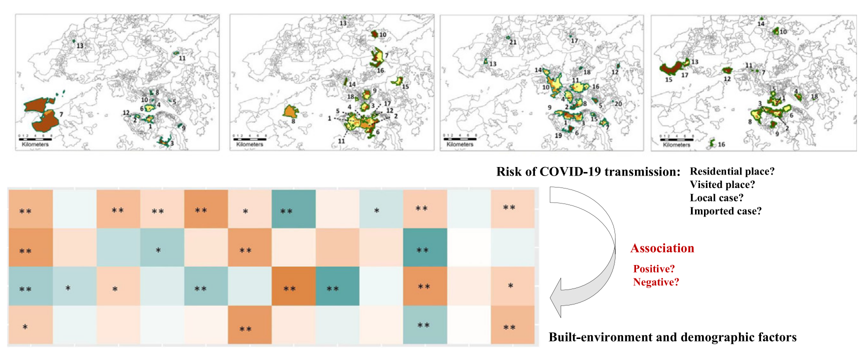
Zihan Kan [a], Mei-Po Kwan [a,b,*], Man Sing Wong [c], Jianwei Huang [a], Dong Liu [d]

[a] Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong, China
[b] Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, Hong Kong, China
[c] Department of Land Surveying and Geo-Informatics, & Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China
[d] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, 1301 W Green St, Urbana, IL 61801, United States

HIGHLIGHTS

- This study identified areas with high COVID-19 transmission risk in Hong Kong.
- Risk was assessed by the number of confirmed cases and the places or venues they visited.
- Built-environment and demographic features that influenced risk were identified.
- The space-time patterns of high-risk locations are complex.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 22 September 2020
Received in revised form 11 January 2021
Accepted 19 January 2021
Available online 27 January 2021

Editor: Wei Huang

Keywords:
COVID-19
Infectious disease
Transmission risk
Built environment
Hong Kong

ABSTRACT

Identifying the space-time patterns of areas with a higher risk of transmission and the associated built environment and demographic characteristics during the COVID-19 pandemic is critical for developing targeted intervention measures in response to the pandemic. This study aims to identify areas with a higher risk of COVID-19 transmission in different periods in Hong Kong and analyze the associated built environment and demographic factors using data of individual confirmed cases. We detect statistically significant space-time clusters of COVID-19 at the Large Street Block Group (LSBG) level in Hong Kong between January 23 and April 14, 2020. Two types of high-risk areas are identified (residences of and places visited by confirmed cases) and two types of cases (imported and local cases) are considered. The demographic and built environment features for the identified high-risk areas are further examined. The results indicate that high transport accessibility, dense and high-rise buildings, a higher density of commercial land and higher land-use mix are associated with a higher risk for places visited by confirmed cases. More green spaces, higher median household income, lower commercial land density are linked to a higher risk for the residences of confirmed cases. The results in this study not only can inform policymakers to improve resource allocation and intervention strategies but also can provide guidance to the public to avoid conducting high-risk activities and visiting high-risk places.

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has seriously threatened global public health since it was first identified in December 2019. Due to its high contagiousness and rapid spread, COVID-19 was

* Corresponding author at: Department of Geography and Resource Management, & Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong, China.
E-mail address: mpk654@gmail.com (M.-P. Kwan).

https://doi.org/10.1016/j.scitotenv.2021.145379

officially declared a pandemic by the World Health Organization (WHO) on March 11, 2020 (World Health Organization, 2020).

During a pandemic, prompt and accurate space-time surveillance of disease are critical for detecting outbreaks and identifying areas with high transmission risks (Lai et al., 2015). As the transmission risk of an infectious disease varies over space and time, monitoring the space-time trends of disease occurrence can highlight the dynamic patterns in risk and help mitigate the spread of the disease. In recent decades, analysis of the spatiotemporal patterns of diseases has become an increasingly common task in the fields of epidemiology, public health and geography (Robertson and Nelson, 2010). The main objectives of analyzing the space-time patterns of a disease are identifying disease clusters, explaining the spatial patterns of the clusters, and predicting the transmission risk of the disease (Caprarelli and Fletcher, 2014).

Existing approaches to exploring the space-time patterns and detecting the high-risk areas of infectious disease include space-time clustering, density estimation and spatial statistics. For instance, local indicators of spatial association (LISA) were used to map clusters and detect hot spots of hand-foot-mouth disease in Beijing, China (Wang et al., 2014). Space-time K function was applied to investigate the space-time interactions and excessive risk of Rift Valley fever transmission in South Africa (Metras et al., 2012). Patterns of dengue cases in the city of Cali, Colombia were mapped with space-time kernel density estimation (Delmelle et al., 2014). Among the various space-time analysis approaches, space-time scan statistics is an effective method for mapping significant clusters of diseases and estimate the associated risk levels based on a variety of statistical models (Kulldorff, 2018). It has been widely used in exploring and mapping the patterns of various diseases, such as vector-borne diseases (Desjardins et al., 2018), rash and respiratory diseases (Takahashi et al., 2008), measles (Tang et al., 2017), dementia (Xu and Wu, 2018), and most recently COVID-19 (Desjardins et al., 2020).

It has long been recognized that spatial context and the built environment contribute to both the initial establishment and dynamic space-time patterns of diseases (Real and Biek, 2007). Research on the relationship between the built environment and spatial distribution of disease can be traced back to the 19th century to the study of the spread of cholera by John Snow (1855). Built environment features at different spatial scales can affect the prevention and spread of infectious diseases. At a smaller scale, as a disease can spread through contaminated objects, certain designs of the physical structure and surface materials of buildings may prevent the spread of infectious diseases. Poor housing conditions and high building density may lead to the problem of inadequate sanitation, which would create an environment conducive to the spread of disease (Pinter-Wollman et al., 2018).

At a larger scale, the built environment affects the space-time patterns of disease transmission by shaping people's activities and social interactions. First, built environment characteristics like the spatial configuration and functional zones of a city significantly affect human mobility and social interactions, which are closely linked to the spread of infectious diseases. For instance, the increasing mobility of people has been identified as the main factor for the emergence of dengue fever (Vazquez-Prokopec et al., 2010). Second, outbreaks of infectious disease are always associated with a disturbance in the usual functioning of public spaces and city infrastructures, which have impacted human mobility and social interactions. In the context of COVID-19, non-pharmaceutical measures including contact tracing and quarantine, social distancing, and the closing of gathering places and venues, have been identified to have a significant impact on human behaviors and social interactions, which are observed to be associated with a reduction in the spread of COVID-19 (Cowling et al., 2020). Kraemer et al. (2020) found that mobility statistics could offer a precise record of the spread of COVID-19 among the cities of China. In the United States, online mapping platforms were developed to provide quantitative information on the changes in people's mobility patterns in response to social distancing guidelines and stay-at-home mandates during the COVID-19 pandemic (Gao et al., 2020; Zhang et al., 2020). As the mobility patterns (e.g., journeys to work) in developed high-density urban societies are highly predictable, transportation infrastructures are usually considered when modeling the transmission of infectious disease (Mei et al., 2015; Mpolya et al., 2014).

There is a vast body of research examining how the built environment shapes people's physical activities, which in turn affect their health outcomes. For instance, characteristics of the built environment such as singular land uses, lower residential densities, poor-quality public open spaces, limited access to public transport, inadequate health care and social service infrastructure are associated with higher risks for major non-communicable diseases, such as being overweight and obesity, physical inactivity and poorer mental health (Koohsari et al., 2013; Garfinkel-Castro et al., 2017; Wang et al., 2020). However, there has been less research on the relationship between the built environment and the transmission risk of infectious disease in space and time. Especially, built environment characteristics associated with high transmission risk of COVID-19 at the local scale have rarely been examined. This study fills this research gap by exploring areas with high COVID-19 transmission risks and the associated built environment and demographic factors in Hong Kong with individual COVID-19 case data. First, this study analyzes the space-time patterns of COVID-19 transmission in Hong Kong using the space-time scan statistic. Second, it examines the built-environment and demographic factors associated with a higher risk of COVID-19 transmission through quartile and correlation analysis. The results reveal much difference in the space-time patterns of high-risk locations as well as the associated factors, which would generate critical insights for developing targeted intervention measures in response to the COVID-19 pandemic.

2. Methods

2.1. Study unit and data

2.1.1. Study area and unit

The study area of this research is Hong Kong, a metropolitan city with a very high population density. As of 2019, 7.4 million residents lived in its 1104 km^2 territory. Large Street Block Group (LSBG), delineated by the Hong Kong Planning Department and used by the census for data reporting purposes, is the study unit in this research. An LSBG is a group of street blocks with similar demographic characteristics. There are 1622 LSBG with 1000 or more residents each in Hong Kong according to the 2016 Hong Kong Census. The study area and study unit are shown in Fig. 1.

2.1.2. COVID-19 case data

The data of daily individual confirmed COVID-19 cases were collected by the Hong Kong Centre for Health Protection and are available to the general public online (at https://data.gov.hk). There were 1013 confirmed cases between January 23 and April 14, 2020. The record of each confirmed case includes the characteristics of each case (e.g., age, gender), type of case (imported, local, close contact with local cases, possibly local, close contact of possibly local cases, and close contact of imported cases), and buildings or venues resided or visited by the confirmed cases in the 14-day period before the day of confirmation. We categorize the case-related locations into residences of the confirmed cases and the locations visited by them (visited locations). For the type of case, imported cases refers to the individuals confirmed with COVID-19 upon returning from abroad, while local cases refer to the individuals who were infected in Hong Kong. We also categorized the individual cases into two general groups: imported cases (including imported and close contact of imported cases) and local cases (including local, close contact with local cases, possibly local, close contact of possibly local cases). The temporal distribution of the confirmed cases in Hong Kong during the study period is shown in Fig. 2, which presents two peaks of confirmed COVID-19 cases.

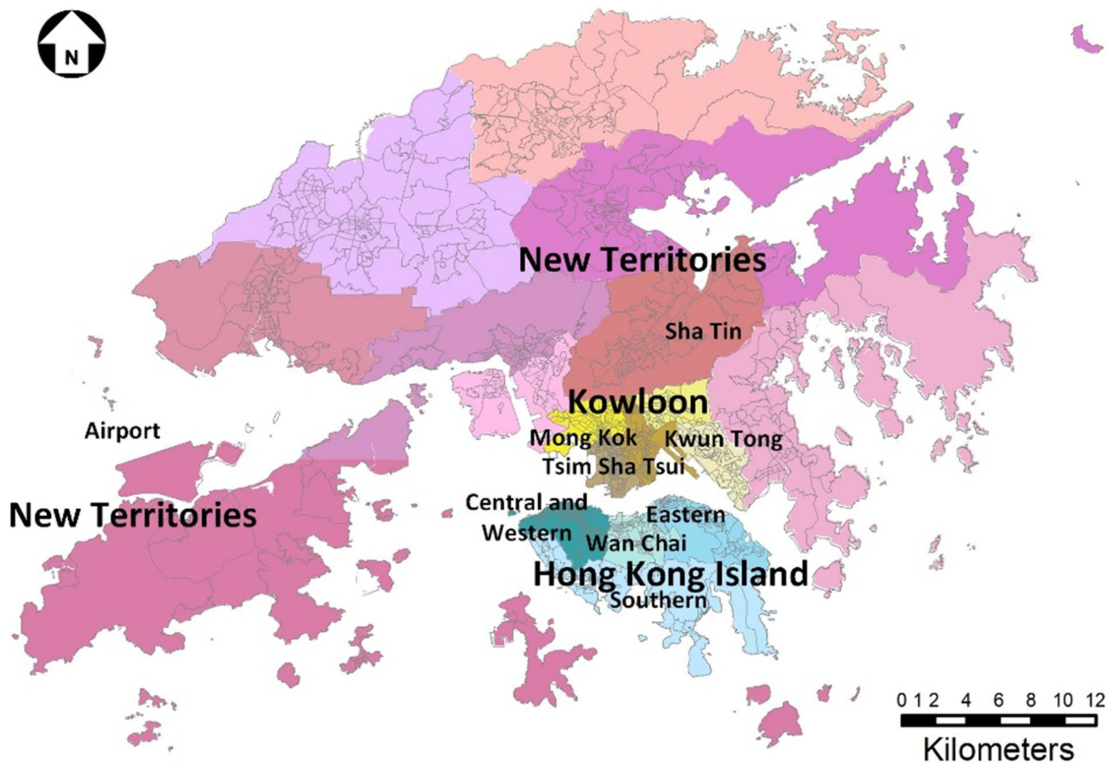


Fig. 1. Study area and study unit.

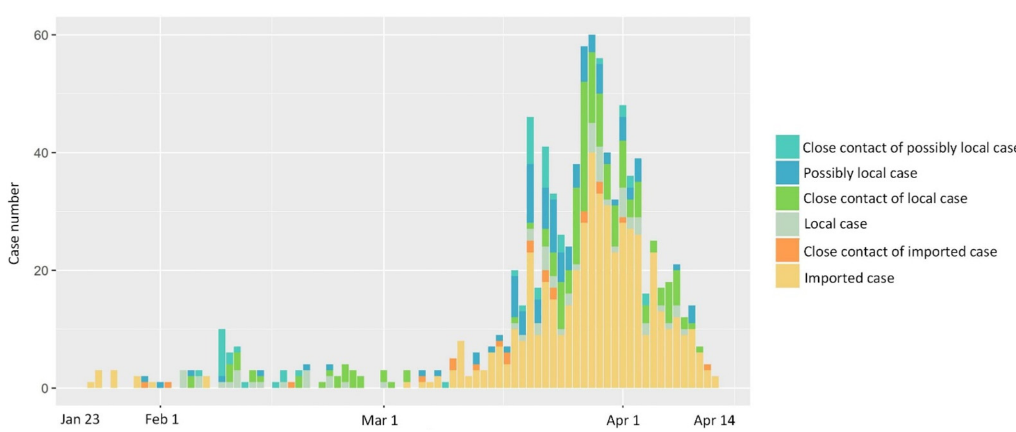


Fig. 2. COVID-19 cases in Hong Kong by dates of confirmation.

Based on the types of buildings/venues and cases, the building/venue-level locations resided or visited by different types of cases are thus categorized into four categories, as Table 1 shows. Residences of the confirmed cases in an LSBG indicate the incidences of COVID-19 in the LSBG, while the locations visited by the confirmed cases in an LSBG reflect the spatial interactions between the confirmed cases and locations in the LSBG, which are potential places of disease transmission.

2.1.3. Demographic and built-environment data

Demographic data used in this study include LSBG-level population and median household income, which are obtained from the Hong Kong Census and Statistics Department. Built environment feature data, covering nodal accessibility, building density, average building height, green spaces, sky view, and land use are calculated from different geospatial data sources.

The nodal accessibility of a transport node, as our first built environment feature, represents how well the node (e.g., a subway or bus station) is connected with other transport nodes in the transport network. As over 90% of the trips in Hong Kong are made by public transport (Hong Kong Transport Department, 2017), this study only considers the public transport networks of the study area, including the Mass Transit Railway (MTR), bus and ferry. Using the transport network data obtained from the Hong Kong Transport Department, the nodal accessibility in each LSBG is calculated based on the connectivity

Table 1
Four categories of building-level locations related to different types of cases.

Location category	Description	Indication
IR	Residences of imported cases	Incidences of COVID-19
LR	Residences of local cases	
IV	Locations visited by imported cases	Intensity of spatial interaction
LV	Locations visited by local cases	

matrix of the nodes of the Hong Kong transportation network in the LSBG. Building density and average building height are derived from the 3D spatial dataset with building geometry and height provided by the Hong Kong Planning Department. The sky view factor is the ratio of the area of sky visible from a given location on the ground to the sky area that is potentially available. It is a 10 m × 10 m raster dataset calculated from multiple data sources including airborne LiDAR data, building GIS data, and land cover data in a previous study (Yang et al., 2015). The area of green spaces in each LSBG is calculated using the Normalized Difference Vegetation Index (NDVI) derived from SPOT-7 Satellite images (2017) with a spatial resolution of 6 m.

Data of a variety of land-use types are acquired from a raster land-use dataset with 27 land-use types and with a spatial resolution of 10 m × 10 m from the Hong Kong Planning Department. The study includes four of these land-use types in its analysis: private residential land density, public residential land density, commercial land density, as well as open spaces and recreation land density. The density of each of these land-use types is obtained by dividing the area of each type of land use in an LSBG by the area of the LSBG. In addition, a land-use mix index (LUMI) is calculated as the degree of the land-use mix for each LSBG based on the notion of entropy, as Eq. (1) shows.

$$LUMI = -\sum_{i=1}^{N} \frac{L_i * \ln L_i}{N} \quad (1)$$

where L_i represents the proportion of the i_{th} type of land use, and N is the total number of land-use types.

2.2. Identifying space-time clusters with high COVID-19 transmission risk

2.2.1. Space-time scan statistic (STSS)

This study analyzes the space-time patterns of COVID-19 in Hong Kong by detecting significant space-time clusters using a space-time scan statistic (STSS) implemented in the SaTScan™ software package (Kulldorff, 2018). We conducted a scan statistic analysis at the LSBG level instead of the individual level because we seek to examine the association between COVID-19 risk and socioeconomic and built environment features, which are measured based on the areal unit of LSBG. This would ensure the consistency of the analytical unit in the study. Given the baseline conditions, an STSS can identify significant space-time clusters of disease locations based on different base models. Relative risk as formulated in Eq. (2) is also calculated for each cluster and each LSBG in a cluster.

$$RR = \frac{n/e}{(N-n)/(N-e)} \quad (2)$$

where RR is the value of the relative risk of an LSBG; n is the total number of confirmed cases in an LSBG; e is the number of expected cases in an LSBG; N is the total number of confirmed cases in the entire study area. Eq. (2) indicates that the relative risk of an LSBG is the estimated risk in the LSBG divided by that outside the LSBG. In this way, a relative risk value higher than 1 means that the LSBG has a higher possibility of exposing to the disease compared with the LSBGs outside it, and the higher the relative risk value, the higher the possibility it would have been exposed to the disease compared with the LSBGs outside it. In the same way, relative risk can also be calculated for a cluster by dividing the estimated risk within a cluster by the estimated risk outside the cluster. A maximum likelihood ratio test is performed to identify the LSBGs with an elevated risk. SaTScan™ uses Monte Carlo simulation to test the statistical significance of the space-time scan statistic. Because the socioeconomic and built environment factors of the study area are at the aggregate LSBG level, by conducting STSS at the LSBG level, we could evaluate the relative risk of each LSBG and analyze its association with the socioeconomic and built environment factors. Further, we use the retrospective SaTScan instead of prospective SaTScan (Hohl et al., 2020; Desjardins et al., 2020) because our study period covers the complete first and second waves of COVID-19 in Hong Kong, and there were no emerging clusters by the end of this period (Fig. 2). In this light, retrospective SaTScan seems more suitable for our dataset.

2.2.2. STSS for residential locations and visited locations

As Section 2.1 mentioned, the building/venue-level locations are categorized into four types: IR, IV, LR, LV. Different types of clusters can be identified based on these four types of locations (e.g., clusters of the residences of local cases or clusters of the locations visited by imported cases). A cluster of residences indicates a higher risk of incidence, while a cluster of visited locations indicates a higher risk of interactions between cases and spatial locations within the cluster, which tend to be conducive to COVID-19 transmission. Therefore, this study identifies space-time clusters of LSBGs and calculates the associated relative risk based on the four types of locations. The discrete Poisson model is chosen as the baseline model in this study because we assume that the number of observations in each LSBG follows a Poisson distribution based on the population or number of buildings/venues in each LSBG. For residential locations (i.e., IR and LR), the null hypothesis is that the expected number of COVID-19 cases is proportional to the population in each LSBG. For visited locations (i.e., IV and LV), however, the null hypothesis is that the expected number of visits is proportional to the number of buildings or venues in each LSBG. Based on the expected number and observed number of COVID-19 cases, the relative risk of COVID-19 transmission for each LSBG or cluster can be calculated by Eq. (2).

2.3. Analyzing the characteristics of clusters with a higher relative risk

This study then uses quartile analysis to depict the trends and patterns of relative COVID-19 risk in each LSBG with respect to the variations in demographic and built environment features. Quartile analysis is similar to decile analysis, which is commonly used as a graphical tool to elucidate the relationships between indicators of different population groups and the associated environmental outcomes (e.g., exposure to air pollution) (Fan et al., 2012). This method sorts each indicator in ascending order and places the corresponding relative risk into each quartile. Because there are variations in the demographic and built-environment indicators in different types of location clusters, this study uses absolute values on the horizontal axis instead of quartile percentages. In this way, the differences in demographic and built environment characteristics between the types of locations can be revealed. Further, Pearson correlation analysis was performed to assess the associations between the relative risk of different location types and different indicators. *t*-Test is conducted to establish whether the correlation coefficient is significant. Using these methods, significant variables contributing to the relative risk of COVID-19 transmission can be identified.

3. Results

3.1. Space-time clustering of COVID-19 cases

Our analysis begins with detecting clusters of different types of locations using the space-time scan statistic. In order to determine the optimum spatial and temporal windows for the space-time statistic, we tested five spatial and temporal windows on the residential data of all confirmed cases in the study period, which are 1%, 5% and 10% of the population at risk as spatial windows, and 10%, 30% and 50% of the study period as temporal windows. We did not select larger spatial and temporal windows in order to avoid detecting extremely large clusters. The results of this exercise using the 5 parameters indicate that the detected clusters are generally stably distributed with different temporal and spatial windows. However, as the spatial and temporal window expands, the spatial and temporal range of the detected clusters becomes larger. We determined that some clusters detected with the spatial windows of 5% of the population at risk and 10% of the population at risk are too large for the study area. For the temporal windows, 10% of the study period failed to detect some clusters that appeared in 30% and 50% of the study period, while the clusters results in 30% and 50% of the study period are similar. The largest time span of the clusters detected by 30% and 50% of the study period are both 23 days, which is within the 30% of the study period. Therefore, we selected 1% of the population at risk and 30% of the study period as the spatial and temporal window. In addition, a cluster is set to have at least 2 confirmed cases to avoid extremely small clusters.

Fig. 3 shows the time distributions for different types of significant clusters (with p values ≤0.05). The horizontal lines in each row represent the detected clusters for each location type (IV, IR, LV and LR). The coverage of each cluster on the horizontal axis depicts the time span of the cluster. The depth of color indicates the number of clusters detected within a period of time. Some important time-variant interventions are also identified and visualized in the form of time line during the study period. As shown in Fig. 3, prior to the restriction on travelers from Mainland China in early February, there are only a few clusters of imported cases from Mainland China. Then in early February, a series of travel restrictions and quarantine measures were implemented aiming at reducing the travel between Hong Kong and Mainland China, including closing all but two border control points and mandatory quarantine for the travelers from Mainland China. The implementation of travel restrictions and quarantine measures between mainland China and Hong Kong corresponds well with a significant drop in the clusters of imported cases as shown in the figure. In comparison, the occurrence of local case clusters persisted throughout February.

Then starting in March, with the deterioration of the pandemic in Europe and other parts of the world (e.g., South Korea), many overseas Hong Kong residents and students returned to Hong Kong, leading to increasing clusters of imported cases from overseas. The Hong Kong government implemented a series of quarantine measures on travelers from many foreign countries or regions from March 1, first on travelers from high-risk areas in Italy or Iran and then from France, Germany, Japan, and the whole of Italy. Clusters of imported cases were at a stable level throughout the first half of March but the clusters of imported cases surpassed that of local cases during this period. From mid-March, the clusters of imported cases started rising sharply due to the deterioration of the pandemic overseas despite the 14-day quarantine requirement being expanded to all travelers arriving at Hong Kong on March 19. As many of these inbound travelers were confirmed with COVID-19 during their quarantine and some returning residents violated the quarantine requirement, the local case clusters also started witnessing a spike.

In late March, the Hong Kong government issued a series of warnings to quarantine breakers, who would face criminal prosecution and compulsory confined quarantine at a government-run quarantine facility. Then on March 29, any group gatherings of more than 4 persons in any public place were banned. In the following days, all karaoke lounges, nightclubs, and mahjong venues were ordered to close. These lockdowns and confinement measures seemed to have helped mitigate the local case clusters, which witnessed a steady decline in April following the lockdowns and confinement measures.

Then after April 8, all inbound travelers arriving at the Hong Kong International Airport are required to submit to COVID-19 testing. The mandatory testing measure appeared to further reduce the number of confirmed cases. Inbound travelers tested positive were immediately sent to the hospital for quarantine and treatment, which minimized the risk of spreading COVID-19. It shows that the series of lockdowns, confinement as well as mandatory testing had effectively reduced both the local and imported case clusters as these measures have, on the one hand, detected the infected individuals upon arrival and minimized the risk of imported cases infecting other locals. On the other hand, these measures reduced the size of group gatherings as well as the number of places for group gatherings.

It can be observed that there were only a small number of clusters from January 23, 2020, the first confirmed COVID-19 case in Hong Kong, to early March 2020. After early March, a large number of imported clusters (both of residences and visited locations) appeared, and many clusters of local cases appeared in mid-March. Therefore, we visualize the spatial distribution of the clusters before and after March 1, 2020, in Figs. 4–7.

Table 2 shows the characteristics of the significant space-time clusters ($p \leq 0.05$) before March 1, 2020, which include the starting and ending times of every cluster, the number of observed cases, the number of expected cases, relative risk of the cluster, p-value and the number of LSBG(s) in each cluster. Since no significant cluster of IV is detected before March 1, 2020, Table 2 only shows clusters of IR, LR and LV. Fig. 4 visualizes the clusters in Table 2, with Fig. 4a showing the clusters within the extent of the entire study area and Fig. 4b showing the enlarged rectangular area in Fig. 4a. The color of each LSBG inside each cluster in Fig. 4 represents the relative risk value of the LSBG.

Table 2 and Fig. 4b show that the first cluster of COVID-19 cases is IR_1, a cluster of the residences of imported cases that appeared during January 29–30 and were located in the private residential area near West Kowloon. This cluster had only one LSBG with a very high relative risk. There were four clusters of local cases, among which LR_1 and LR_3 appeared during February 9–10. LR_1 was located in Wan Chai District with a relative risk value of 164.1. LR_2 and LR_3, both with higher relative risk values, were located in the Eastern District and composed of only one LSBG. LR_4 was located in Kwun Tong with only one LSBG and high relative risk value. According to the news report, clusters LR_1-LR_4 were family case clusters. Before March, a cluster of visited places (i.e., LV_1) appeared during February 16–27 in Eastern District, which included 29 LSBGs with a relative risk value of 81.4. Buildings and venues in this cluster were visited 17 times by confirmed cases during this period.

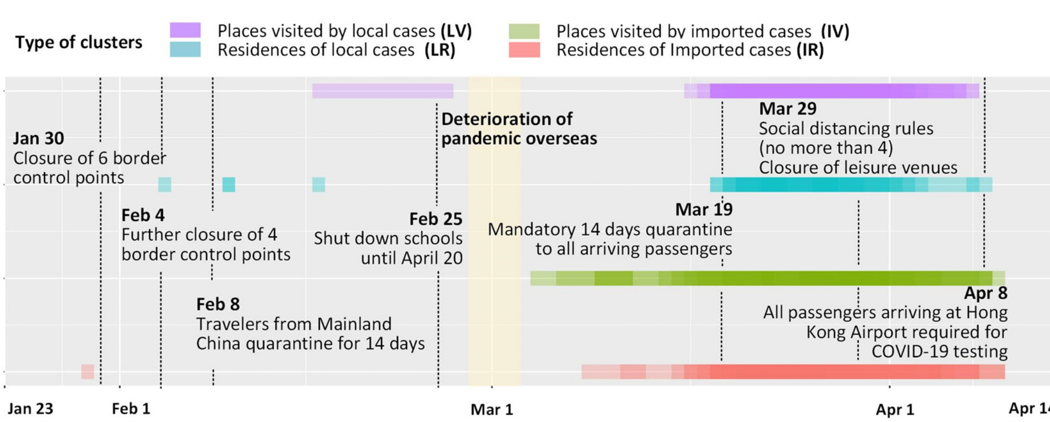


Fig. 3. The temporal distributions of significant clusters of different types of cases.

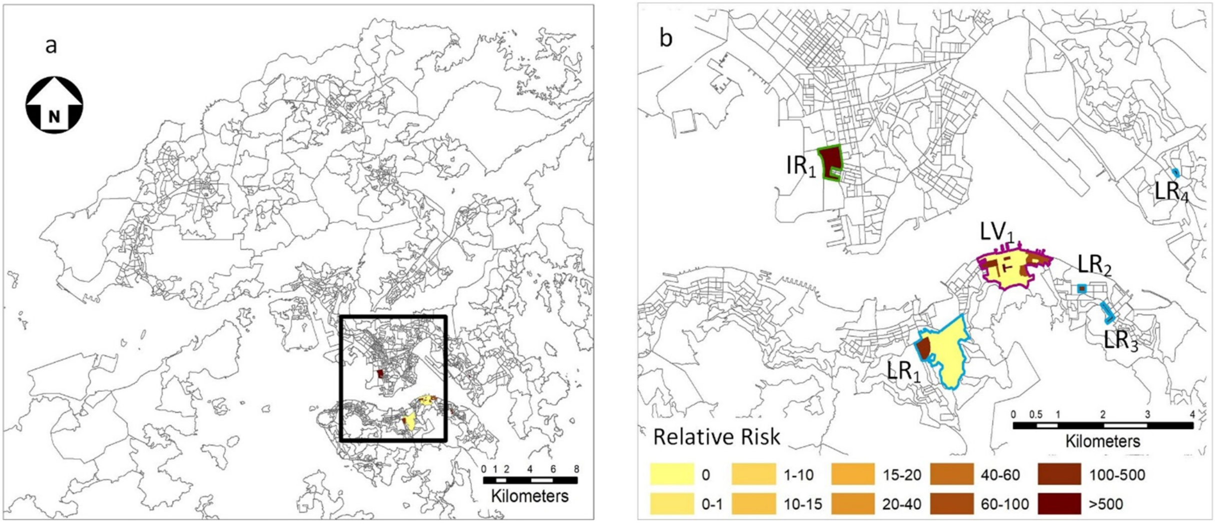


Fig. 4. Spatial distribution of space-time clusters of COVID-19 in Hong Kong before March 1, 2020.

Visited places in this cluster include a private clinic, a public hospital, a cafeteria and a Buddhist temple, where a group outbreak occurred.

Tables 3, 4 and Fig. 5a and b show the characteristics and spatial distribution of the space-time clusters of cases (LR and IR) after March 1, 2020. The cluster IDs in Fig. 5a and b correspond to the cluster IDs in Tables 3 and 4. In Table 3 and Fig. 5a, a total of 67 LSBGs in Clusters 1–3, 9 and 12 were located in Wan Chai District, Central and Western District, East District and Southern District on Hong Kong Island, with relative risk values from 55.5 to over 500. There are 36 LSBGs in Clusters 4–6, 8 and 10 located in Kowloon, in which the largest cluster is Cluster 6 (with 32 LSBGs and a relative risk of 19.7) in Tsim Sha Tsui, which has a high residential density. Only 10 LSBGs in Clusters 7, 11 and 13 were located in the New Territories, with a relative risk of 26.6, 47.0 and 100. Among them, Cluster 7 was located near the airport, where there are many residential buildings for the staff of the Hong Kong International Airport and airlines who are high-risk groups for COVID-19 infection.

Table 4 and Fig. 5b show the space-time clusters of the residences of the imported cases (IR) after March 1, 2020. It can be seen by comparing Fig. 5a with Fig. 5b that the spatial extent of the IR cluster is larger than that of the LR clusters. Table 4 shows that most of the significant clusters occur in later March and early April. During this period, a large number of overseas students and residents returned to Hong Kong due to the increasingly severe COVID-19 outbreak in certain countries abroad (e.g., the U.K.). There were 18 significant IR clusters in total, out of which 10 clusters (Clusters 1, 2, 4–6, 11,12,17) with 126 LSBGs were located on Hong Kong Island. The high-risk areas include residential areas in Wan Chai District, Eastern District and Southern District, as well as residential neighborhoods at Mid-Levels, which is an affluent residential area in Central and Western District on Hong Kong Island. Clusters 3, 9, and 18 were located in Kowloon Peninsula, with 62 LSBGs. The largest cluster in Kowloon is Cluster 3 in Tsim Sha Tsui, with 32 LSBGs and a relative risk of 16.2. Different from the LR clusters, there are many IR clusters in the New Territories (i.e., Clusters 7, 8, 10, 13–16), with 15 LSBGs.

Tables 5, 6 and Fig. 6a and b show the space-time clusters of visited buildings or venues by local cases (LV) and imported cases (IV). As shown in Table 5 and Fig. 6a, there were 7 LV clusters (Clusters 1, 3, 6–7, 9, 15 and 19) on Hong Kong Island with a total of 165 LSBGs and 241 visits. Cluster 1, with a relative risk of 140.1, received 162 visits during March 18–April 8. This cluster was located near Central on Hong Kong Island, which has a high density of dining and entertainment venues where multiple group outbreaks occurred (e.g., a night-club and gym cluster in Lan Kwai Fong). Six clusters (Clusters 2, 4, 5, 8, 11, 16) with 238 LSBGs and 139 visits were located in Kowloon. Cluster 2, with a relative risk of 500.0, was located in Tsim Sha Tsui and received 70 visits during the week between March 25 and April 2. A group outbreak took place in a Karaoke bar in this area. Cluster 3 was located in Wan Chai District on Hong Kong Island and received 35 visits by

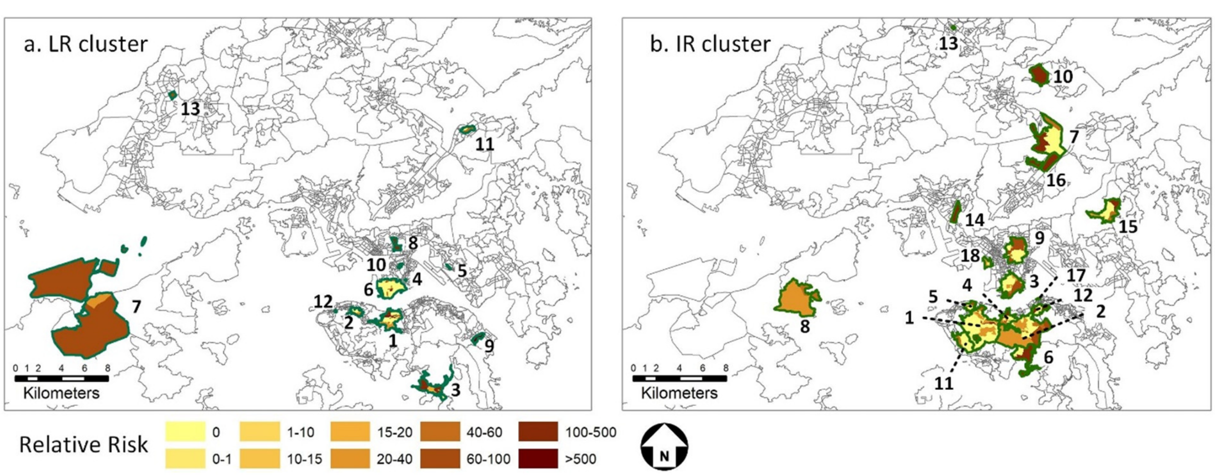


Fig. 5. Space-time clusters of the residential buildings of local and imported cases.

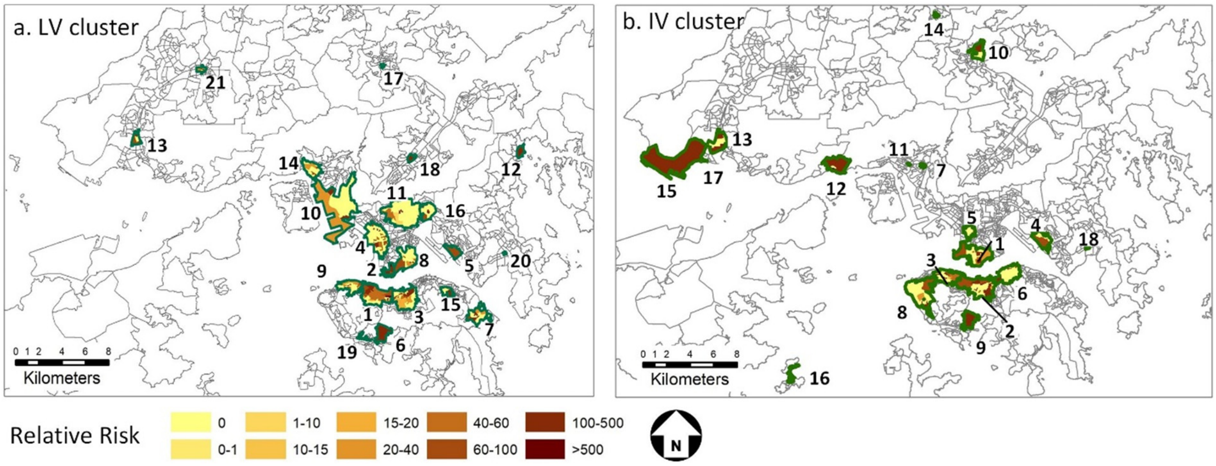


Fig. 6. Space-time clusters of places visited by local and imported COVID-19 cases.

confirmed cases during March 16–April 7. This is a busy commercial area with a high traffic volume. Although there were more clusters (8 clusters) in the New Territories than on Hong Kong Island and in Kowloon, the clusters (Cluster 9, 10, 12–14, 17, 18, 20, 21) were smaller and with fewer visits compared with the other two areas.

Table 6 and Fig. 6b show the clusters of buildings/venues visited by the imported cases (IV) of COVID-19. The Hong Kong Government required that individuals returning to Hong Kong after March 19 must go through a 14-day mandatory quarantine. Some of these visited buildings (e.g., bars) were visited by returnees before March 19 while some other buildings were hotels used by some of the returnees to self-quarantine. For the imported cases, although there were not many close contacts of imported cases in Fig. 2, all locations resided or visited by imported cases should be deemed as at-risk since the imported cases may unknowingly infect other people in those places. Besides, since COVID-19 can be transmitted through surface contamination (Ong et al., 2020; Razzini et al., 2020; Santarpia et al., 2020), it is possible that places resided or visited by imported cases are capable of infecting others via contaminated surfaces. Some local cases whose sources of infection are not clear might have been infected in places visited by imported cases. By knowing the built environment characteristics associated with the relative risk of imported cases, the government can implement preemptive sanitizing and deep-cleaning measures in those places in order to prevent transmission at those high-risk places. Moreover, the government can also mandate strict social distancing measures (e.g., installation of plexiglass) at places associated with a high risk of imported cases to minimize the risk of imported cases unconsciously infecting others. A total of 95 visits in Clusters 2, 3, 6, 8 and 9 involving 159 LSBGs were located on Hong Kong Island. These places have high commercial densities and received visits from confirmed cased at a public hospital, a private clinic and several entertainment venues. 70 visits in Clusters 1, 4, 5 involving 86 LSBGs were located in Kowloon. In the New Territories, there were 10 clusters (Clusters 7, 10–18) with 31 LSBGs and 36 visits. It can be seen that clusters in the New Territories are large in number, small in size and sparse in distribution but have a high relative risk. This is because the population and building densities in the New Territories are lower than those on Hong Kong Island and in Kowloon, which makes the expected number of visits by confirmed cases relatively low. As a result, the relative risks of the detected clusters would be higher compared with areas with a higher population or building densities.

3.2. Association between demographic features, built environment features and the relative risk of COVID-19

We use both quartile and correlation analysis to explore the association between the relative risk of COVID-19 transmission and relevant demographic and built environment features. The demographic characteristics of an LSBG included in this analysis are median household income and population density. Built-environment features include nodal accessibility, building density, building height, green spaces (derived using the NDVI), the sky view factor, private residential land density, public residential land density, commercial land density, open spaces and recreation land density and the land-use mix index (LUMI). The results of the quartile analysis are shown in Fig. 7. In each graph in the figure, each line represents the trend of relative risk (RR) with the increase of quartiles of each feature of the clustered LSBGs (there are four types of clusters: IR. IV, LR and LV). We also include the quartile distribution of each feature of all LSBGs as a reference to reflect the distribution of each feature of the clustered LSBGs. Using this figure, we can investigate the demographic and built-environment features associated with both the occurrence of clusters and a higher relative risk of LSBGs within the clusters.

The figure shows that the clustered LSBGs have higher median household income, most of which exceed the median household income of HKD 25,000 for Hong Kong in 2016 (HKCSD, 2016). Especially, the median household income of the LSBGs in IR clusters is higher compared with that of all LSBGs and it far exceeds those of the LSBGs of other types of clusters. It indicates that the imported cases have higher income levels and median household income can characterize the occurrence of the IR clusters. This is probably because most of the imported cases are people working or studying overseas, who may have higher socioeconomic status than the average person. It can also be observed in the figure that the population density of the LSBGs in IR clusters and LR clusters is lower compared with that of all LSBGs, and lower than LV clusters and IV clusters. This is perhaps due to the similar fact indicated by the distribution of median household income: that is, the imported cases with higher socioeconomic status also tend to live in areas with lower population density.

In terms of built-environment features, Fig. 7 shows that compared with all LSBGs, IR and LR clusters are associated with lower building density, lower building height, fewer green spaces, higher private residential land density, lower public residential density and lower open and recreation land density. IR clusters also have less sky view compared with that of all LSBGs. There is no obvious difference between LR and IR clusters and that of all LSBGs. The reason for less sky view and fewer green spaces in IR and LR clusters than that in all LSBGs might be because there are many areas in the countryside (e.g., country parks) with more green spaces and better sky views in Hong Kong. Not many people live in those areas and the number of cases in those areas is too low to form a cluster. However, Fig. 7 also shows that the higher relative risk of LR and IR clusters is associated with more green spaces, which is probably because the

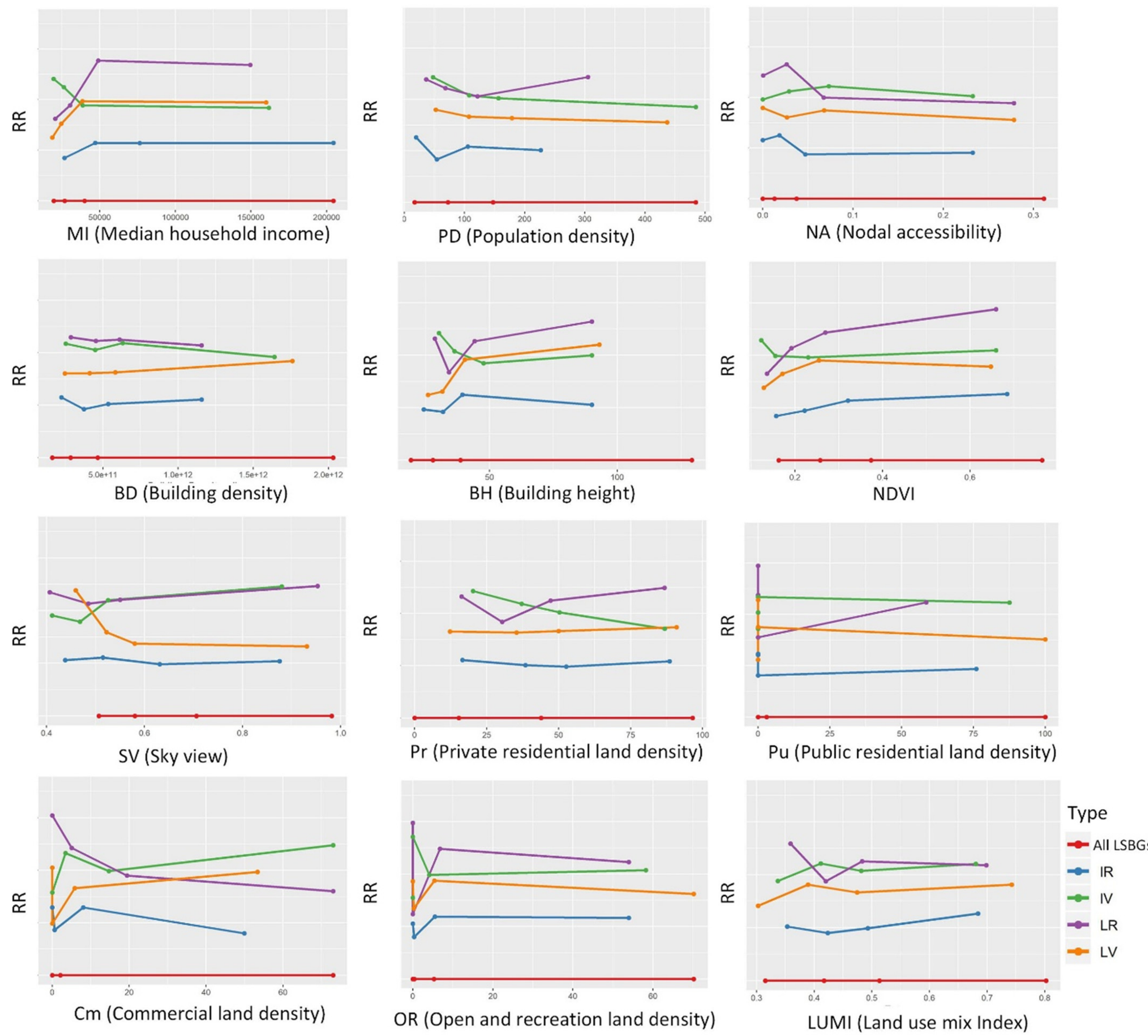


Fig. 7. Change in relative risk with respect to interquartile range increase of different factors in each group.

lower population density in the LR and IR clustered LSBGs leads to higher relative risk levels.

Fig. 7 shows that the building height of IV and LV clusters are also lower than that for all LSBGs. Private residential land density of IV and LV clusters are higher compared with that of all LSBGs. But within the IV clusters, a higher risk is associated with lower building height, lower private residential land density and more sky view. The LV clusters with a higher building density, higher building height and less sky view tend to have a higher risk level. It can also be observed that the building density in the high-risk areas visited by the confirmed cases (LV and IV) is higher than that in high-risk residence areas (IR and LR). There is no obvious association between the relative risk of all

Table 2
Space-time clusters of COVID-19 in Hong Kong before March 2020.

Cluster	Duration	Observed	Expected	Relative risk	*p*-Value	No. of LSBG
IR$_1$	Jan 29–Jan 30	3	0.0023	>500	<0.001	1
LR$_1$	Feb 9–Feb 10	4	0.25	164.1	0.007	8
LR$_2$	Feb 16–Feb 17	3	0.0056	500.0	0.012	1
LR$_3$	Feb 9–Feb 10	3	0.0043	>500	0.006	1
LR$_4$	Feb 4–Feb 5	3	0.003	>500	0.002	1
LV$_1$	Feb 16–Feb 27	17	0.21	81.4	<0.001	29

Table 3
Space-time clusters of local COVID-19 cases (LR) in Hong Kong after March 2020.

Cluster	Duration	Observed	Expected	Relative risk	*p*-Value	No. of LSBG
1	Mar 22–Apr 04	34	0.67	55.5	<0.001	39
2	Mar 18–Mar 30	22	0.39	59.9	<0.001	19
3	Mar 20–Apr 02	11	0.12	91.1	<0.001	6
4	Apr 01–Apr 09	6	0.01	450.4	<0.001	1
5	Mar 27–Mar 28	4	<0.01	>500	<0.001	1
6	Mar 19–Apr 01	12	0.63	19.7	<0.001	32
7	Mar 21–Apr 01	10	0.38	26.6	<0.001	2
8	Mar 19–Mar 24	4	0.01	500.0	<0.001	1
9	Apr 07–Apr 08	3	<0.01	>500	0.002	1
10	Mar 26–Mar 27	3	<0.01	>500	0.002	1
11	Mar 18–Mar 23	6	0.13	47.0	0.004	7
12	Mar 23–Apr 03	4	0.03	127.7	0.02	2
13	Mar 20–Mar 21	4	0.04	100.0	0.043	1

Table 4
Space-time clusters of imported COVID-19 cases (IR) in Hong Kong after March 2020.

Cluster	Duration	Observed	Expected	Relative risk	*p*-Value	No. of LSBG
1	Mar 18–Apr 08	44	1.26	38.2	<0.001	30
2	Mar 18–Apr 06	23	1.08	22.3	<0.001	30
3	Mar 16–Apr 05	19	1.22	16.2	<0.001	32
4	Mar 17–Apr 08	16	0.90	18.3	<0.001	27
5	Mar 20–Apr 08	15	0.79	19.6	<0.001	21
6	Mar 19–Mar 20	5	0.01	450.1	<0.001	4
7	Mar 16–Apr 03	9	0.20	45.0	<0.001	4
8	Mar 22–Apr 10	9	0.21	44.4	<0.001	2
9	Mar 20–Apr 10	14	0.98	14.71	<0.001	26
10	Mar 22–Mar 23	4	0.01	>500	<0.001	1
11	Mar 08–Mar 27	9	0.36	25.4	<0.001	4
12	Mar 20–Mar 23	6	0.09	66.5	<0.001	9
13	Apr 04–Apr 05	3	<0.01	>500	0.001	1
14	Mar 11–Mar 13	4	0.02	186.8	0.004	1
15	Mar 15–Apr 05	6	0.16	36.8	0.012	5
16	Mar 18–Mar 22	3	0.01	396.5	0.021	1
17	Mar 22–Mar 30	4	0.04	103.3	0.032	1
18	Mar 18–Apr 10	7	0.35	20.1	0.045	4

Table 6
Space-time clusters of visited buildings of imported COVID-19 cases (IV) in Hong Kong after March 2020.

Cluster	Duration	Observed	Expected	Relative risk	*p*-Value	No. of LSBG
1	Mar 17–Apr 09	46	0.62	88.2	<0.001	40
2	Mar 14–Apr 05	43	0.82	61.5	<0.001	46
3	Mar 18–Apr 10	28	0.80	38.8	<0.001	47
4	Mar 11–Apr 01	12	0.26	48.3	<0.001	20
5	Mar 11–Apr 03	12	0.26	48.1	<0.001	26
6	Mar 22–Apr 09	11	0.38	29.7	<0.001	55
7	Mar 25–Apr 09	5	0.01	503.8	<0.001	1
8	Mar 04–Mar 12	8	0.15	54.3	<0.001	10
9	Mar 22–Mar 30	5	0.02	245.0	<0.001	1
10	Mar 10–Mar 11	4	0.01	466.0	<0.001	10
11	Mar 28–Apr 08	3	0.00	>1000	<0.001	1
12	Mar 10–Mar 11	4	0.01	273.3	<0.001	3
13	Mar 16–Mar 17	4	0.02	253.2	0.001	11
14	Mar 15–Mar 22	3	0.00	967.4	0.002	1
15	Mar 09–Mar 10	4	0.02	199.5	0.002	1
16	Mar 06–Mar 11	5	0.08	66.1	0.007	1
17	Mar 09–Mar 10	2	0.00	>1000	0.012	1
18	Mar 19–Mar 20	2	<0.01	>1000	0.03	1

types of clusters and public residential land density and open spaces and recreation land density. In comparison, an increase in commercial land density is associated with an increase in the relative risk of IV clusters and LV clusters. The figure shows that the relative risk for LSBGs in LV, LR and IR clusters tends to increase with an increase in green spaces (NDVI). This may be due to that the locations with a higher COVID-19 risk tend to be wealthier places with more green spaces resided by confirmed cases. Similarly, an increase in land-use mix is also associated with an increase in relative risk for LV, IV and IR clusters. It indicates that commercial land density and land-use mix are important factors that influence the relative risk assessed by both the number of confirmed cases and the number of venues visited by confirmed cases.

Fig. 8 further visualizes the correlation between the relative risk of the LSBGs in different types of clusters and different demographic and built environment features. It also shows the significant contributing factors of the relative risk assessed by the number of confirmed cases or the number of places visited by confirmed cases. In general, the correlation between the relative risk of LSBGs and different features confirms the association trends in Fig. 7. When relative risk is assessed by the number of local confirmed cases (LR), areas with a higher risk of COVID-19 have higher median household income, more green spaces, lower building density and lower commercial land density. When relative risk is assessed by the number of imported confirmed cases (IR), areas with higher COVID-19 risk have higher median household income, higher land-use mix, more green spaces and lower commercial land density. In comparison, when relative risk is assessed by the number of venues visited by the local confirmed cases (LV), areas with a high risk of COVID-19 have higher median household income, higher nodal accessibility, higher building density, higher building height, more green spaces, less sky view, lower public residential land density and higher land-use mix. It indicates that the places with higher transportation accessibility, denser and more high-rise buildings, and less sky view tend to attract people to visit and conduct activities and thus increase the risk of COVID-19 transmission. Those areas usually have a high density of commercial and entertainment venues as shown in Fig. 6a. When relative risk is assessed by the number of venues visited by the imported cases (IV), areas with higher relative risk have lower median household income, lower population density, higher nodal accessibility, lower building height, more sky view, lower private residential land density, higher commercial land density and higher land-use mix. This indicates that a higher commercial land use density and a higher land-use mix potentially increase the relative risk due to more visits and social activities of confirmed cases.

Table 5
Space-time clusters of buildings or venues visited by local COVID-19 cases (LV) in Hong Kong after March 2020.

Cluster	Duration	Observed	Expected	Relative risk	P-value	No. of LSBG
1	Mar 18–Apr 08	162	1.56	140.1	<0.001	32
2	Mar 25–Apr 02	70	0.16	500.0	<0.001	9
3	Mar 16–Apr 07	35	1.29	28.7	<0.001	45
4	Mar 18–Mar 31	25	1.02	25.4	<0.001	121
5	Mar 20–Mar 31	12	0.15	80.8	<0.001	4
6	Mar 17–Apr 02	10	0.08	124.2	<0.001	1
7	Mar 23–Apr 08	12	0.31	39.0	<0.001	19
8	Mar 18–Apr 01	13	0.60	22.2	<0.001	47
9	Mar 18–Mar 31	13	0.75	17.6	<0.001	50
10	Mar 20–Mar 27	10	0.41	25.0	<0.001	16
11	Mar 18–Apr 04	14	1.28	11.2	<0.001	45
12	Mar 29–Apr 02	5	0.03	166.9	<0.001	1
13	Mar 18–Mar 29	6	0.08	75.1	<0.001	7
14	Mar 20–Apr 03	8	0.30	27.1	<0.001	17
15	Mar 22–Mar 27	5	0.05	101.6	0.001	13
16	Apr 01–Apr 03	5	0.06	88.9	0.002	12
17	Mar 24–Mar 25	3	<0.01	>500	0.003	2
18	Mar 27–Mar 28	3	0.01	>500	0.008	2
19	Mar 27–Apr 01	4	0.03	139.4	0.011	5
20	Mar 18–Mar 21	5	0.08	62.3	0.011	1
21	Mar 28–Apr 06	3	0.01	390.4	0.018	1

4. Discussion

This study explored the areas with high COVID-19 risk in Hong Kong and the associated built environment and demographic characteristics. Temporal and spatial clustering is usually present in the distribution of both communicable and chronic diseases. Clusters of a communicable disease in certain places indicate a higher risk of disease transmission. The clustering patterns of a communicable disease are related to certain built environment and socioeconomic characteristics. For instance, a communicable disease like COVID-19 spreads mainly through close contact between carriers and other people and people's contact with surfaces of objects with the virus (Ong et al., 2020; Razzini et al., 2020; Santarpia et al., 2020). Thus, excessive transmission risk may be a result of certain built-environment features (e.g., the density of high-risk venues like pubs and restaurants as well as transport facilities like public transit stations) that attract more people to visit, conduct their daily activities or engage in social interactions (as observed in Huang et al. (2020) concerning the COVID-19 pandemic in Hong Kong).

The space-time clustering results in Section 3.1 highlights the time periods and spatial locations with high transmission risk in Hong Kong during the first two waves of the COVID-19 pandemic in the city. By categorizing case-related locations into four different types — residences of

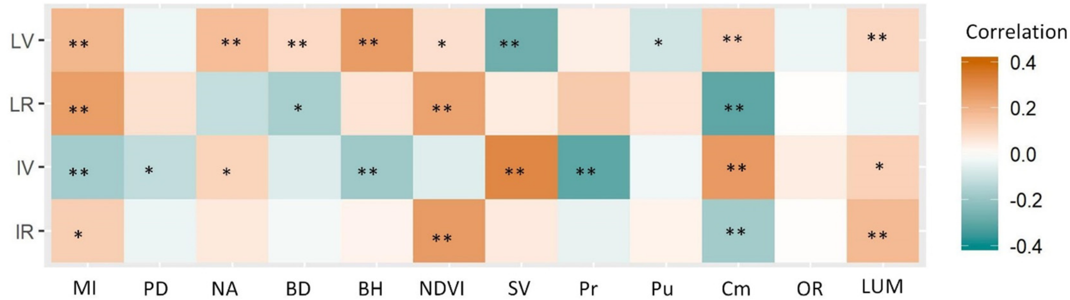


Fig. 8. Correlation of the relative risk with the demographic and built-environment features (*: $p < 0.1$; **: $p < 0.05$). MI: median household income; PD: population density; NA: nodal accessibility; BD: building density; BH: building height; GS: green spaces; SV: sky view; Pr: private residential land density; Pu: public residential land density; Cm: commercial land density; OR: open spaces and recreation land density; LUM: land-use mix. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

imported cases (IR), residences of local cases (LR), buildings/venues visited by imported cases (IV) and buildings/venues visited by local cases (LV) — the study generated more detailed knowledge on the space-time patterns of COVID-19 risk in Hong Kong. In this study, a cluster of residences or a cluster of locations or venues visited by infected persons means the clustering pattern of the locations are associated with certain risk factors rather than just because there are more population or venues than other places. Note that we also evaluated the relative risk of each cluster and each LSBG in a cluster by taking into account the number of expected cases in them. The clusters with a higher relative risk can thus be considered as having a higher risk than the normal. A cluster of residential places indicates that there is a higher risk of disease incidence in the cluster, and people living there are also at a higher risk of being infected. A cluster of visited places indicates a higher intensity of social interactions between humans associated with high-risk places and venues like bars and pubs, which also increase the risk of disease transmission. It is possible that a cluster in a residential area is caused by cases from the same family. However, it will lead to a higher risk of disease incidence and people living in the area are also at a higher risk of being infected than people living at other locations.

Since the first COVID-19 case was imported from Mainland China on January 23, 2020, the Hong Kong Government took a series of measures to mitigate COVID-19 transmission, including closing all but two border checkpoints with Mainland China and essentially denying all but Hong Kong residents' entry into Hong Kong, and closing certain venues such as schools and social distancing. The situation of the COVID-19 pandemic in Hong Kong before March 2020 was stable: the number of confirmed cases remained low, especially for local cases. However, as the COVID-19 pandemic overseas became more serious in early March, many Hong Kong citizens and students residing overseas began to return to Hong Kong. The increasing number of imported cases led to an increasing number of local cases, which is confirmed by that the emergence of the clusters of local cases lagged behind that of the imported cases in Fig. 3. In terms of the spatial patterns of the high-risk areas, Figs. 4–6 show that there are significant differences between different types of clusters. As indicated by the residences of the imported and local cases, the local cases and imported cases seemed to involve different population groups as people who work or study overseas may have higher socioeconomic status when compared to those of the local cases, which may result in different patterns of high-risk areas between local and imported cases. For instance, Fig. 5 shows that local cases are more likely to reside in high-density areas on Hong Kong Island and Kowloon, while the residences of the imported cases distribute more widely, including very affluent areas like Mid-Levels.

According to the contact tracing information available via the government webpage on COVID-19, many of the high-risk residential buildings are due to family members being cross-infected in short order. These cases of family-related outbreaks were usually caused by one or more family members who were infected when visiting certain places or venues. By focusing on specific case-related locations (i.e., residences and places visited), high-risk areas of COVID-19 transmission as well as high-risk activities are revealed in this study. The high-risk areas in Fig. 6 indicate that pubs, fitness centers, and restaurants on Hong Kong Island and in Kowloon are high-risk places, which are likely to be visited by local cases.

The analysis of the association and correlation between demographic and built environment features and relative risks further indicates the key factors shaping the COVID-19 landscape in Hong Kong. In the cluster detection, the null hypothesis is that the expected number of COVID-19 cases is proportional to the population in each LSBG. A cluster means that there is excessive risk of COVID-19 instead of the occurrence of the disease. Therefore, we explored the demographic and built-environment features associated with higher risks of COVID-19 based on detecting the space-time clusters. Figs. 7 and 8 also indicate that the factors linked to higher risks of different types of clusters (locations of residence and visited places) are different. First, median household income has a positive correlation with the relative risk of residential locations of both local and imported cases. This means that areas with higher household income tend to have a higher COVID-19 risk, which is contrary to what was reported in much of the literature: many studies in other countries or cities reported negative correlations between economic factors (income level or GDP) and the transmission risk of infectious diseases where people with higher income tend to have healthier lifestyles and better access to healthy residential environmental and healthcare resources (National Academies of Sciences, 2018). For instance, in the United States, 40% of low-income people were identified to have a higher risk of infecting COVID-19, compared to 24% of higher-income people (Raifman and Raifman, 2020). Actually, there are few reported cases in poor neighborhoods in Hong Kong during the time period of this study. As the activity spaces of low-income people are more confined than those of the high-income people in Hong Kong (Tao et al., 2020), and the activity spaces of these two population groups may not overlap significantly (Wang and Li, 2016), low-income people may have lower possibilities of visiting the high-risk locations and thus are exposed to lower infection risk.

A higher population density is usually considered as an important factor associated with a higher risk of infectious disease transmission because high population density means higher chances of contact between people (Xu et al., 2019). However, for the COVID-19 pandemic in Hong Kong, population density has only a weak negative correlation with the locations visited by imported cases and does not have a significant correlation with other types of locations. This is because some suburban areas in the New Territories with low population density are identified as high-risk clusters. The higher risk of these suburban areas, with higher proportions of their areas being recreation spaces like country parks, may be related to higher levels of certain kinds of human activities (e.g., hiking and picnicking) after the COVID-19 outbreak as reported in the news.

The correlation between the risks of different location types and built environment features would further characterize different types of locations with high transmission risk of COVID-19. For the locations visited by confirmed cases (IV and LV), the high-risk locations visited by imported cases have higher transportation accessibility, lower building height, better sky view, higher proportions of commercial land and higher land-use mix. While the high-risk locations visited by local cases are places with higher transportation accessibility, denser and more high-rise buildings, less sky view, higher commercial land density and higher land-use mix. Figs. 7 and 8 indicate that visited places with higher COVID-19 risk have both high transportation accessibility and land-use diversity. Surprisingly, green space is found to have a positive correlation with the COVID-19 risk of both residential and visited locations. This is perhaps because the locations with a higher risk tend to be wealthier places during the study period, and those places usually have higher proportions of green spaces. Also, suburban areas with country parks tend to attract more visits during the pandemic (e.g., several bank employees were found hiking in a country park while they were supposed to be working at home) and thus also tend to have a higher COVID-19 risk. The correlation analysis between relative COVID-19 risk and demographic/built-environment features in this study generates knowledge on how different features of places affect the transmission risk of COVID-19. The patterns of risky areas associated with the confirmed cases (especially for the local cases) may provide a useful reference for understanding the transmission risk of other infectious diseases with a similar transmission mechanism as COVID-19. It can also provide evidence that helps target interventions to the places with high COVID-19 risk and places where the residents tend to have a higher risk of being infected.

This study contributes to advancing knowledge and policy-making in important ways. First, our knowledge about the impact of the built and social environments on the spread of COVID-19 is still highly limited to date. This study contributes to the literature by analyzing the space-time dynamics of areas with high COVID-19 transmission risk and the associated built environment and demographic features in Hong Kong using individual COVID-19 case data. It is the first study to distinguish space-time clusters of COVID-19 cases as imported and local and COVID-19-related locations as residential and visited, which provides more refined and meaningful findings about areas with high COVID-19 transmission risk and better inform the development of more effective control measures.

With regard to policy implications, the results of our study can inform policymakers to enhance the effectiveness of intervention measures by developing more targeted strategies to reduce people's exposure to the risk factors of COVID-19 transmission. On one hand, the locations of high transmission risk (e.g., pubs, karaoke venues and restaurants) identified in the study can inform the formulation and implementation of stricter mitigation measures such as banning indoor activities in those places. On the other hand, the identified areas with high transmission risk and areas with certain built environment features associated with a higher COVID-19 transmission risk can be used to inform people to avoid conducting high-risk activities and visiting these high-risk places. Further, certain features of the built environment can be modified or dynamically managed in order to promote healthy behaviors and reduce the risk of COVID-19 transmission.

In this study, both the residential and visited places of each confirmed case were obtained, which makes it possible to analyze the case-related COVID-19 patterns and their relationships with various demographic and built environment features. It may not be possible to replicate the study in other countries and cities with large numbers of confirmed cases every day, where it is less feasible for contact and activity tracing to generate the detailed individual data used in this study, which is a limitation of this research. However, the results of this study still provide a useful picture of the areas with a high COVID-19 risk in a high-density and socioeconomically segregated city like Hong Kong. Moreover, we aggregated individual confirmed cases at the LSBG level so that statistical modeling can be used to test the underlying reasons for the clustering of the cases. As aggregating individual data into areal units may raise problems including the modifiable areal unit problem and the instability of values for small population areas, further studies may consider applying the kernel density method to take the full advantage of individual level information (Shi, 2009, 2010). Another limitation of this study is that the dates of confirmation are used in our analysis which tend to lag behind the exact time of being infected and the dates of symptom onset.

5. Conclusion

This study identifies the high-risk areas of COVID-19 in Hong Kong between January 23 and April 14, 2020, and examines the characteristics of different types of high-risk locations by analyzing the association between the relative risk and various demographic and built environment features for each high-risk cluster. The results reveal much difference in the space-time patterns of high-risk locations as well as the contributing factors, which provide useful knowledge for better understanding the COVID-19 pandemic in Hong Kong and improving intervention strategies. Future work includes further examining the activity patterns of different population groups and their impact on the transmission of infectious disease in different social contexts. The environmental features associated with disease incidence will also be explored in our future study.

CRediT authorship contribution statement

Z.K., M.P.K. and J.H. conceived and designed the study. Z.K. collected the data and conducted data analysis. M.S.W., J. H., and D.L. helped process the data. Z.K. and M.P.K. interpreted the results. Z.K. and M.P.K wrote the original draft. Z.K., M.P.K., M.S.W., J.H. and D.L. revised the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by a grant from the Research Committee on Research Sustainability of Major RGC Funding Scheme of the Chinese University of Hong Kong. Zihan Kan was supported by an RGC Postdoctoral Fellowship from the Research Grants Council of Hong Kong (PDFS2021-4S08). Man Sing Wong thanks the funding support from a grant by the General Research Fund (Grant no. 15602619). Dong Liu was supported by a Marion G. Russell Graduate Fellowship from the University of Illinois at Urbana-Champaign.

References

Caprarelli, G., Fletcher, S., 2014. A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. Parasitology 141, 581–601.

Cowling, B.J., Ali, S.T., Ng, T.W., Tsang, T.K., Li, J.C., Fong, M.W., Liao, Q., Kwan, M.Y., Lee, S.L., Chiu, S.S., 2020. Impact Assessment of Non-pharmaceutical Interventions Against COVID-19 and Influenza in Hong Kong: An Observational Study (MedRxiv).

Delmelle, E., Dony, C., Casas, I., Jia, M., Tang, W., 2014. Visualizing the impact of space-time uncertainties on dengue fever patterns. Int. J. Geogr. Inf. Sci. 28, 1107–1127.

Desjardins, M.R., Whiteman, A., Casas, I., Delmelle, E., 2018. Space-time clusters and co-occurrence of chikungunya and dengue fever in Colombia from 2015 to 2016. Acta Trop. 185, 77–85.

Desjardins, M., Hohl, A. & Delmelle, E. 2020. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. Applied Geography, 102202.

Fan, X., Lam, K.-C., Yu, Q., 2012. Differential exposure of the urban population to vehicular air pollution in Hong Kong. Sci. Total Environ. 426, 211–219.
Gao, S., Rao, J., Kang, Y., Liang, Y. & Kruse, J. 2020. Mapping county-level mobility pattern changes in the United States in response to COVID-19. Available at SSRN 3570145.
Garfinkel-Castro, A., Kim, K., Hamidi, S., Ewing, R., 2017. Obesity and the built environment at different urban scales: examining the literature. Nutr. Rev. 75, 51–61.
Hohl, A., Delmelle, E.M., Desjardins, M.R., Lan, Y., 2020. Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. Spatial and Spatio-temporal Epidemiology 34, 100354.
Hong Kong Census and Statistics Department (HKCSD), 2016. Hong Kong 2016 Population By-census - Thematic Report:Household Income Distribution in Hong Kong.
Hong Kong Transport Department, 2017. Public Transport Strategy Study: Final Report. Hong Kong Transport Department, Hong Kong.
Huang, J., Kwan, M.-P., Kan, Z., Wong, M.S., Kwok, C.Y.T., Yu, X., 2020. Investigating the relationship between the built environment and relative risk of COVID-19 in Hong Kong. ISPRS Int. J. Geo Inf. 9 (11), 624.
Koohsari, M.J., Badland, H., Giles-Corti, B., 2013. (Re) Designing the built environment to support physical activity: bringing public health back into urban design and planning. Cities 35, 294–298.
Kraemer, M.U., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., Du Plessis, L., Faria, N.R., Li, R., Hanage, W.P., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368, 493–497.
Kulldorff, M., 2018. SaTScan User Guide 2018.
Lai, P.C., Chow, C.B., Wong, H.T., Kwong, K.H., Kwan, Y.W., Liu, S.H., Tong, W.K., Cheung, W.K., Wong, W.L., 2015. An early warning system for detecting H1N1 disease outbreak–a spatio-temporal approach. Int. J. Geogr. Inf. Sci. 29 (7), 1251–1268.
Mei, S., Chen, B., Zhu, Y., Lees, M.H., Boukhanovsky, A., Sloot, P.M., 2015. Simulating city-level airborne infectious diseases. Comput. Environ. Urban. Syst. 51, 97–105.
Metras, R., Porphyre, T., Pfeiffer, D.U., Kemp, A., Thompson, P.N., Collins, L.M., White, R.G., 2012. Exploratory space-time analyses of Rift Valley fever in South Africa in 2008–2011. PLoS Negl. Trop. Dis. 6, e1808.
Mpolya, E.A., Yashima, K., Ohtsuki, H., Sasaki, A., 2014. Epidemic dynamics of a vector-borne disease on a villages-and-city star network with commuters. J. Theor. Biol. 343, 120–126.
National Academies of Sciences, Engineering, and Medicine, 2018. In: Ogawa, V.A., Shah, C.M., Nicholson, A. (Eds.), Urbanization and Slums: Infectious Diseases in the Built Environment: Proceedings of a Workshop. The National Academies Press, Washington, DC (148 p).
Ong, S.W.X., Tan, Y.K., Chia, P.Y., Lee, T.H., Ng, O.T., Wong, M.S.Y., Marimuthu, K., 2020. Air, surface environmental, and personal protective equipment contamination by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from a symptomatic patient. Jama 323 (16), 1610–1612.
Pinter-Wollman, N., Jelić, A., Wells, N.M., 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, 20170245.
Raifman, M.A., Raifman, J.R., 2020. Disparities in the population at risk of severe illness from covid-19 by race/ethnicity and income. Am. J. Prev. Med. 59 (1), 137–139.
Razzini, K., Castrica, M., Menchetti, L., Maggi, L., Negroni, L., Orfeo, N.V., Pizzoccheri, A., Stocco, M., Muttini, S., Balzaretti, C.M., 2020. SARS-CoV-2 RNA detection in the air and on surfaces in the COVID-19 ward of a hospital in Milan, Italy. Sci. Total Environ. 742, 140540.
Real, L.A., Biek, R., 2007. Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. J. R. Soc. Interface 4, 935–948.
Robertson, C., Nelson, T.A., 2010. Review of software for space-time disease surveillance. Int. J. Health Geogr. 9, 16.
Santarpia, J.L., Rivera, D.N., Herrera, V.L., Morwitzer, M.J., Creager, H.M., Santarpia, G.W., Crown, K.K., Brett-Major, D.M., Schnaubelt, E.R., Broadhurst, M.J., Lawler, J.V., 2020. Aerosol and surface contamination of SARS-CoV-2 observed in quarantine and isolation care. Sci. Rep. 10 (1), 1–8.
Shi, X., 2009. A geocomputational process for characterizing the spatial pattern of lung cancer incidence in New Hampshire. Ann. Assoc. Am. Geogr. 99 (3), 521–533.
Shi, X., 2010. Selection of bandwidth type and adjustment side in kernel density estimation over inhomogeneous backgrounds. Int. J. Geogr. Inf. Sci. 24 (5), 643–660.
Snow, J., 1855. On the Mode of Communication of Cholera (John Churchill).
Takahashi, K., Kulldorff, M., Tango, T., Yih, K., 2008. A flexibly shaped space-time scan statistic for disease outbreak detection and monitoring. Int. J. Health Geogr. 7, 14.
Tang, X., Geater, A., Mcneil, E., Deng, Q., Dong, A., Zhong, G., 2017. Spatial, temporal and spatio-temporal clusters of measles incidence at the county level in Guangxi, China during 2004–2014: flexibly shaped scan statistics. BMC Infect. Dis. 17, 243.
Tao, S., He, S.Y., Kwan, M.-P., Luo, S., 2020. Does low income translate into lower mobility? An investigation of activity space in Hong Kong between 2002 and 2011. J. Transp. Geogr. 82, 102583.
Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., Ritchie, S.A., 2010. Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. PLoS Negl. Trop. Dis. 4, e920.
Wang, D., Li, F., 2016. Daily activity space and exposure: a comparative study of Hong Kong's public and private housing residents' segregation in daily life. Cities 59, 148–155.
Wang, J., Cao, Z., Zeng, D.D., Wang, Q., Wang, X., Qian, H., 2014. Epidemiological analysis, detection, and comparison of space-time patterns of Beijing hand-foot-mouth disease (2008–2012). PLoS One 9, e92745.
Wang, P., Goggins, W.B., Zhang, X., Ren, C., Lau, K.K.-L., 2020. Association of urban built environment and socioeconomic factors with suicide mortality in high-density cities: a case study of Hong Kong. Sci. Total Environ. 139877.
World Health Organization. 2020. WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020.
Xu, W., Wu, C., 2018. Detecting spatiotemporal clusters of dementia mortality in the United States, 2000–2010. Spatial and spatio-temporal epidemiology 27, 11–20.
Xu, B., Tian, H., Sabel, C.E., Xu, B., 2019. Impacts of road traffic network and socioeconomic factors on the diffusion of 2009 pandemic influenza A (H1N1) in Mainland China. Int. J. Environ. Res. Public Health 16, 1223.
Yang, J., Wong, M.S., Menenti, M., Nichol, J., 2015. Modeling the effective emissivity of the urban canopy using sky view factor. ISPRS J. Photogramm. Remote Sens. 105, 211–219.
Zhang, L., Ghader, S., Pack, M. L., Xiong, C., Darzi, A., Yang, M., Sun, Q., Kabiri, A. & Hu, S. 2020. An interactive COVID-19 mobility impact and social distancing analysis platform. medRxiv.