Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Spatio-temporal patterns of the COVID-19 pandemic, and place-based influential factors at the neighborhood scale in Tehran

Azadeh Lak a, Ayyoob Sharifi b, Siamak Badr c, Alireza Zali d, Ali Maher e, Ehsan Mostafavi f, Davood Khalili g

a Department of Planning and Urban Design, Faculty of Architecture and Urban Planning, Shahid Beheshti University, Tehran, Iran
b Hiroshima University, Graduate School of Humanities and Social Science & Network for Education and Research on Peace and Sustainability (NERPS), Japan
c Faculty of Architecture and Urban Planning, Shahid Beheshti University, Tehran, Iran
d Functional Neurosurgery Research Center, Shohada Tajrish Neurosurgical Center of Excellence, Shahid Beheshti University of Medical Sciences, Tehran, Iran
e School of Management and Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran
f Department of Epidemiology and Biostatistics, Research Centre for Emerging and Reemerging Infectious Diseases, Pasteur Institute of Iran, Tehran, Iran
g Prevention of Metabolic Disorders Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
COVID-19
Spatio-temporal analysis
Demographic factors
Place-based factors
Geographically weighted regression (GWR)
Tehran

ABSTRACT

Since its emergence in late 2019, the COVID-19 pandemic has attracted the attention of researchers in various fields, including urban planning and design. However, the spreading patterns of the disease in cities are still not clear. Historically, preventing and controlling pandemics in cities has always been challenging due to various factors such as higher population density, higher mobility of people, and higher contact frequency. To shed more light on the spread patterns of the pandemic, in this study we analyze 43,000 confirmed COVID-19 cases at the neighborhood level in Tehran, the capital of Iran. To examine spatio-temporal patterns and place-based factors contributing to the spread of the pandemic, we used exploratory spatial data analysis and spatial regression. We developed a geo-referenced database composed of 12 quantitative place-based variables related to physical attributes, land use and public transportation facilities, and demographic status. We also used the geographically weighted regression model for the local examination of spatial non-stationarity. According to the results, population density ($r^2 = 0.88$) and distribution of neighborhood centers ($r^2 = 0.59$), drugstores ($r^2 = 0.64$), and chain stores ($r^2 = 0.59$) are the main factors contributing to the spread of the disease. Additionally, density of public transportation facilities showed a varying degree of contribution. Overall, our findings suggest that demographic composition and major neighborhood-level physical attributes are important factors explaining high rates of infection and mortality. Results contribute to gaining a better understanding of the role of place-based attributes that may contribute to the spread of the pandemic and can inform actions aimed at achieving Sustainable Development Goals, particularly Goals 3 and 11.

1. Introduction

The new coronavirus disease (COVID-19) has become a universal health concern due to its high transmission rate and its rapid spread throughout the world. In fact, the United Nations has defined the pandemic as a human, social, and economic crisis. Although the socio-economic consequences are particularly evident in developing countries, developed countries have also been hardly hit by the pandemic. With the spread of the pandemic, some countries like Iran have been compelled to take emergency measures such as closing city borders, suspending schools, stopping production lines, and restricting the movement of population, which have caused great harm to economic development as well as public health.

Since the beginning of the pandemic, researchers in many parts of the world have shown interest in understanding its driving mechanisms and spatio-temporal spread patterns, particularly in meso- and micro-urban scales (Hu, Roberts, Azevedo, & Milner, 2020; Kim & Bostwick, 2020; Li, Peng, He, Wang, & Feng, 2020). Among other things, these
studies have aimed at developing mechanisms for the prevention and control of the pandemic by identifying risk hotspots, and highlighting risk factors that may contribute to its transmission in cities (Li et al., 2020; Mansour, Al Kindi, Al-Said, Al-Said, & Atkinson, 2021; Sharifi & Khavarian-Garmsir, 2020). Hui et al. (2020) compared the transmission rate of the COVID-19 pandemic with its precedents and argued that a prolonged pandemic is likely to pose severe threats to public health. They suggested that, across different scales, an integrated approach to the human-environment-animals health is needed to ensure well-being and sustainability of communities. Using the exponential growth and maximum likelihood estimation method, Liu, Hu et al. (2020) determined the transmission mechanisms of COVID-19 in Wuhan and discovered that the average incubation period was 4.8 days, and the primary regeneration index reached 2.90 (95% Confidence Interval (CI): 2.32–3.63) and 2.92 (95% CI: 2.28–3.67). Zhonghun, Xing and Xue (2020) examined the spatial-temporal features of the spread of COVID-19 in Guangdong province, concluding that the adopted prevention and control measures were effective and high-risk areas were mostly located in economically developed parts of the province. Similarly, Liu et al. (2020) conducted a statistical analysis to examine the spatio-temporal characteristics of the transmission path of the pandemic in Zhuhai, China. They found that connections to the epicenter area as well as family gatherings were the leading causes of the spread. Likewise, according to Xie et al. (2020), the inflow of population from Wuhan and strong economic connections were the fundamental causes of the disease spread. Other factors included population distribution pattern, average temperature, transport accessibility, and quality of medical facilities.

The effects of socio-economic factors have also been examined in many studies. In a comparison of patient hotspots in New York City (NYC) and Chicago, Maroko, Nash, and Pavilion (2020) demonstrated that hotspots in NYC include the communities of the working- and middle-income classes, whose wages are above poverty levels. Similarly, in Chicago hotspot areas are located in low-income neighborhoods of the city that are characterized by remarkably high rates of unemployment and poverty, as well as, large numbers of Black residents. Taking a multi-dimensional approach, Mollalo, Vahedi, and Rivera (2020) conducted a GIS-based spatial modelling to measure the effects of socio-economic, environmental, behavioral, demographic, and topographic factors on the incidence of COVID-19 in the USA. Their findings show that various explanatory variables such as median household income, income inequality, the proportion of black females, and the proportion of nurse practitioners affect the spatial distribution of COVID-19 patients. Adopting a similar nationwide approach, Kang, Choi, Kim and Choi (2020) explored the spatio-temporal patterns of the disease and measured the provincial spatial associations in China between January and February 2020. The results indicated that the confirmed cases had significant spatial characteristics. These results were in congruence with those of Hu et al. (2013) concerning the spread of the SARS epidemic. In a more detailed analysis of different contributing factors, You, Wu, and Guo (2020) demonstrated that the morbidity rate of COVID-19 in Wuhan on 22 February 2020 was positively correlated with population density, construction land area proportion, total retail sales of consumer goods per unit of land area, value-added of tertiary industry per unit of land area, density of public green space, and density of aged population; and negatively correlated with the average building size, GDP per unit of land area, and density of hospitals. Similar multivariable studies have also been conducted elsewhere. For example, a study based on a set of environmental variables, Almagro and Orme-Hutchinson (2020) developed a regression model to evaluate the statistical significance of the relationship between neighborhood measures (i.e., population density, commuting patterns, and health insurance controls) and response variables (i.e., demographic status and incidence of COVID-19) in the neighborhoods of NYC. They found out that occupation was an important factor explaining the observed COVID-19 patterns. Indeed, people with more social interactions turned out to be more vulnerable to the pandemic. Several other studies in other parts of the world have also examined the relationship between COVID-19 infection and mortality rates and explanatory variables like neighborhood characteristics (Borjas, 2020), age structure (Kulu & Dorey, 2020), social and economic bonds (Sannigrahi, Pilla, Basu, Basu, & Molter, 2020), and physical and socio-economic characteristics (Khavarian-Garmsir, Sharifi, & Moradpour, 2021). This indicates that multiple factors may be influential and should be considered.

The literature discussed above has improved our understanding of the spread patterns and dynamics of the pandemic. However, there are still gaps and shortcomings that need to be addressed. A major gap is that existing studies have mainly focused on the city or regional scales, and there are few studies at the neighborhood level (Kim & Bostwick, 2020; Maroko et al., 2020). While city- and regional-level analyses can provide useful insights into the overall spread patterns and dynamics of the disease and the effectiveness of response measures and policies (e.g., partial, or blanket lockdowns), they fall short of revealing detailed intra-city dynamics and variations. Conducting neighborhood-based analysis can contribute to addressing this gap.

Another noteworthy gap is that existing research has mainly used correlation and regression analysis methods, and newer spatial analysis methods have largely been neglected (Wang et al., 2020), Xiong, Wang, Chen, and Zhu (2020) argued that ‘time-series spatial statistics and factor analysis’ can help discover the spatio-temporal features of the pandemic and inform scientific decision-making towards its control and prevention (Xiong et al., 2020). With this in mind, Geographic Information System (GIS) seems to be an effective tool for investigating the spatial distribution of infectious diseases and developing solutions to combat them. In fact, GIS has turned into a useful tool for analyzing and visualizing the spread of COVID-19 (Li et al., 2020; Mansour et al., 2021). The spread of the disease in a city depends on the rate of human mobility (Li et al., 2020) and may be affected by demographic and the place-based attributes of urban neighborhoods. Accordingly, in the present study, after identification of the hotspots of confirmed patients in Tehran- the capital city of Iran-, we utilize a Geographically Weighted Regression model (GWR) to specify how they could explain the spatio-temporal differences in the spread of the disease based on some quantitative factors related to land use, demographic characteristics, and physical and public transportation facilities. Tehran, as the case study, is a national hub of administrative, cultural, and economic activities. It is the capital and most populous city of the country and covers a large geographical area. The city has been the epicenter of the pandemic in Iran and is one of the most hardly hit cities in the region. These characteristics make it a suitable case for studying the spatio-temporal patterns of the pandemic.

The study aims to (i) understand the spatio-temporal evolution and patterns of the disease spread at the neighborhood level and identify the hotspots, and (ii) explore the quantitative place-based factors contributing to the spread of the pandemic at the neighborhood scale. Addressing these questions can improve our understanding of influential factors related to human behavior in urban physical settings, thereby informing planners and authorities of necessary prevention and control measures that should be taken to mitigate the impacts of the pandemic. To the best of our knowledge, this is the first attempt to model local geographic distribution of the COVID-19 pandemic at the neighborhood level, and on a city-wide scale. Neighborhood-level analysis is needed as it allows more detailed and granular analysis of the spread patterns and associated socio-economic and physical factors. We hypothesize that the spread of the COVID-19 pandemic is likely to have significant associations with place-based variables at the meso- and micro-scales. The other scientific contributions of the study are as follows. First, a new time-series geo-database for Tehran is created for the first five months of the pandemic (20 February – 20 July), which offers information on the spatio-temporal patterns of the spread of the virus. Focus on Tehran is important as it is one of the hardest hit cities in the Middle East that has not been well studied in the literature. Then, place-based indicators...
(physical attributes, transportation facilities, and land use) and demographic status are studied to investigate the potential factors that contribute to the spread of the pandemic at the neighborhood level through GWR methods. The results can offer urban health and management authorities with useful information on necessary measures that need to be taken for controlling the pandemic at meso- and micro-levels. This can enable them to develop sustainable and inclusive plans for dealing with health risks, thereby contributing to the achievement of Sustainable Development Goals, particularly Goals 3 and 11. Finally, the methodological approach adopted in this study can also be used in other cities, especially developing country cities, to better understand the spread patterns of the pandemic and develop appropriate context-specific mitigation and response measures.

2. Methods and materials

2.1. Study design and data

The scale of the study is the municipal neighborhoods of Tehran, the capital city of Iran. The city has an area of more than 700 km² and consists of 22 municipal districts and 354 administrative neighborhoods (Fig. 1). The population density of the neighborhoods ranges between 40 and 580 persons per hectare (Fig. 2). According to the 2018 statistics of Tehran municipality, the total population of the city is 8,938,686 (4,472,268 females and 4,466,825 males). This population consists of 3,041,133 households with an average of 2.9 persons per household. In Districts 1-7, the population is generally of higher socio-economic status (Ghaaze, Rafiejian, & Dadashpoor, 2020).

Our cross-sectional study was carried out based on data collected in those medical centers of Tehran, which are allocated to the treatment of COVID-19 patients (Gholamzadeh et al., 2020). Soon after the identification of the first several confirmed cases of COVID-19 in Iran, the Ministry of Health and Medical Education started to record the patients’ information in two databases, namely, Medical Care Monitoring Center (MCMC) and Hospitals’ Information Management (HIM). These databases are updated on a daily basis and are officially used for gathering, analyzing, and reporting the COVID-19 data in Tehran (Gholamzadeh et al., 2020).

All patients that have been examined with a Polymerase Chain Reaction (PCR) test and admitted and hospitalized as confirmed COVID-19 cases between 20 February and 20 July 2020 were studied. February 20 is the date when Iran reported its first COVID-19 case and July 20 is when official data for the first five months of the pandemic were released. In addition, according to later data updates, July 20 is the peak date of the second wave of the pandemic in Tehran (Jalali et al., 2020). In addition to cases confirmed by laboratory tests, patients with clinical symptoms on chest Computed Tomography (CT) scan findings who initially had negative test results were also diagnosed as highly suspected COVID-19 patients and hospitalized. We reviewed the data retrieved from the databases as mentioned earlier based on the patients’ national ID, personal data, date of admission and discharge, the beginning date of the symptoms, and the patients’ home address. Finally, after screening the data and removing duplicates, we analyzed the geographical location of the COVID patients. The locations (i.e., addresses) of all inpatients and outpatients were analyzed after geocoding in GIS. The Research Ethics Committee of SBMU, Tehran, Iran (IR.SBMU.RETECH.REC.1399.070) approved the study.

2.2. Selection of variables

A recent review of COVID-19 research on cities indicates that the spatial spread of the disease has been affected by multiple socio-economic and environmental factors (Sharifi & Khavarian-Garmsir, 2020). As mentioned earlier, there is still a need for a more comprehensive and context-specific exploration of aspects like transmissibility, means and modes of transmission, as well as the risk factors that may contribute to the spread of the disease (Ren et al., 2020; Wong, Leo, & Tan, 2020). This is, particularly, warranted at the neighborhood level since recent epidemiological studies have indicated that person-to-person transmission is more common at this level during the incubation period of the disease (Chan et al., 2020; Huang, Xia, Chen, Shan, & Wu, 2020; Yu, Zhu, Zhang, & Han, 2020). In this study, we mainly considered significant factors such as quantitative place-based variables as essential factors that may contribute to the spread of the disease. We selected demographic and urban form indicators in order to examine how they may affect the COVID-19 transmission. Significance of these indicators has been demonstrated in previous research (Hu et al., 2020; Lak, Asl, & Maher, 2020; Mansour et al., 2021; Sharifi & Khavarian-Garmsir, 2020). From a demographic perspective, population density and the percentage of population aged above 65 are two major indicators of the level of vulnerability to the disease (Franch-Pardo, Napoletano, Rosete-Verges, & Billa, 2020; Kiahgahi, Rifi, & Liaw, 2020; Sharifi & Khavarian-Garmsir, 2020). Population density is argued to be a likely risk factor for the spread of the pandemic. This is because maintaining social distancing is likely to be more challenging in high-density areas, thereby increasing the risk of exposure to the virus (Sharifi & Khavarian-Garmsir, 2020). Also, population aged above 65 are demonstrated to be more vulnerable to infectious diseases due to the characteristics of their immune system (Mansour et al., 2021). As with the urban form, ten quantitative variables were gathered and classified into three categories: 1) physical setting factors, namely, building density and the ratio of green and open spaces, to represent the level of openness of the neighborhood. These are important indicators as they may influence the degree of compliance with social distancing rules and regulations. In fact, dense areas with limited open and green spaces may increase the risk of exposure to COVID-19 (Megahed & Ghoneim, 2020; Sharifi & Khavarian-Garmsir, 2020); 2) land uses and utilities to meet the daily needs of the residents. These include availability and distribution pattern of land uses and utilities such as neighborhood centers and chain stores that allow meeting daily socio-economic, mental, and physical needs of citizens while complying with social distancing measures. In addition, medical facilities (i.e., drugstores, radiology centers, and selected hospitals for COVID-19 patients) are needed to ensure appropriate response to the pandemic (Franch-Pardo et al., 2020; Ren et al., 2020); and 3) transportation facilities, the distribution of which is demonstrated to affect the transmission patterns of the pandemic (i.e., subway stations, bus stops, and gas stations) (Ren et al., 2020). These are often frequently visited locations and, therefore, may contribute to the spread of the pandemic.

These items were all indexed according to their density, at a 450 m × 450 m resolution, at the neighborhoods of Tehran. While there is no universal consensus on the definition of neighborhood, this scale is argued to be reasonable as it allows meeting daily needs within walkable distances (e.g., according to New Urbanist2 thinkers). Data related to the above variables were retrieved from the comprehensive plan of Tehran municipality (Municipality, 2020). Next, Geographically Weighted Regression (GWR) was utilized to analyze the relationship between the dependent variable (COVID-19 incidence rate) and potential explanatory variables.

2.3. Spatial analysis

2.3.1. Getis-Ord Gi* statistics

The possibility of analyzing multiple layers in a single map is a major advantage of the GIS. Various types of georeferenced and non-georeferenced data are combined in GIS (Alemdar, Kaya, & Čodur, 2020). Accordingly, using GIS for spatial analysis of urban settlement attributes allows planners and policy makers to take better-informed decisions towards creating safer and healthier cities (Alemdar et al., 2020; Kaya, Alemdar, & Billa, 2020; Kiaghadi, Rifai, & Liaw, 2020); and Sharifi & Khavarian-Garmsir, 2020). Population density is argued to be a likely risk factor for the spread of the pandemic. This is because maintaining social distancing is likely to be more challenging in high-density areas, thereby increasing the risk of exposure to the virus (Sharifi & Khavarian-Garmsir, 2020). Also, population aged above 65 are demonstrated to be more vulnerable to infectious diseases due to the characteristics of their immune system (Mansour et al., 2021). As with

---

2 https://www.cnu.org/resources/what-new-urbanism
Fig. 1. The 22 districts of Tehran.

Fig. 2. Population density in Tehran’s neighborhoods.
Getis-Ord \( G^* \) is a local statistics technique in GIS to determine new locations with significant hot and coldspot clusters. This technique can be used to identify disease clusters within the neighborhoods and analyze context-specific conditions (Sharifi, Mesgar, Kasraee, & Jahangiri-rad, 2020). The collected COVID-19 data was analyzed to identify disease hotspots in the neighborhoods. Equations 1, 2, and 3 are used to calculate the local Getis-Ord \( G^* \) statistics. According to Panahi et al. (2020) and Tewara et al. (2019), \( X_j \) is the attribute value for feature \( j \); \( w_{ij} \) is the spatial weight between \( i \) and \( j \); and \( n \) is equal to the total number of features' attributes.

\[
G^*_i = \frac{\sum_{j=1}^{n} w_{ij} X_j - \sum_{j=1}^{n} \sum_{k=1}^{n} w_{ik} X_k}{\sqrt{\sum_{j=1}^{n} w_{ij}^2 - \sum_{j=1}^{n} \sum_{k=1}^{n} w_{ik} w_{jk}}} \\
\overline{G} = \frac{\sum_{i=1}^{n} G_i}{n} \\
S = \sqrt{\frac{\sum_{i=1}^{n} G_i^2}{n} - \overline{G}^2}
\]

### 2.3.2. GWR spatial regression analysis

We used regression analysis to examine the spatial relationship between the distribution of COVID-19 hotspots and place-based variables. Several studies have demonstrated the utility of the local regression models (GWR, MGWR) for explaining the spatial variation in COVID-19 infection rate in urban and regional levels (Han et al., 2021; Li, Ma, & Zhang, 2021; Maiti et al., 2021; Mansour et al., 2021; Sannigrahi et al., 2020). By using the GWR local regression model, we can obtain an Eq. (2) for each element of the dependent variable dataset to capture the geographic variations (Tewara et al., 2019). Based on the kernel-weighted regression, GWR is highlighted as an extension to general regression models to relax the spatial correlation among variables and allow for "parameters to vary spatially" (Brunsdon, Fotheringham, & Charlton, 1996). Instead of estimating global values for regression parameters, GWR derives the local factors for each distinctive location, thus incorporating the geographic context (Mollalo et al., 2020; Oshan, Smith, & Fotheringham, 2020). GWR is defined by Eq. (4) (Fotheringham & Oshan, 2016):

\[
y_i = \beta_0 + \sum_{j=1}^{m} \beta_j X_j + \epsilon_i, i = 1, 2, ..., n
\]

In Eq. (4), in city \( i \), \( y_i \) is the COVID-19 spread rate, \( \beta_0 \) is the intercept, \( \beta_j \) is the \( f_j \) regression parameter, \( X_j \) is the value of the \( f_j \) explanatory factors, and \( \epsilon_i \) is a random error term. According to Fotheringham and Oshan (2016), for each explanatory variable and in each city, parameter estimates are given by:

\[
\hat{\beta}(i) = (X' W(i) X)^{-1} X' W(i) y
\]

Where, according to Mollalo, Mao, Rashidi, and Glass (2019), \( ^t \) is the vector of parameter estimates \((m \times 1)\), \( X \) is the matrix of the selected explanatory variables \((n \times m)\), \( W(i) \) is the matrix of spatial weights \((n \times n)\), and \( y \) is the vector of observations of the dependent variable \((m \times 1)^t\). \( W(i) \) is a diagonal matrix that is constructed by the weights of observations based on their distance from location \( i \), and is calibrated using a locally weighted regression (Fotheringham & Oshan, 2016; Mollalo et al., 2020). A kernel function and a bandwidth should be specified to calculate \( W(i) \). The most widely used kernel functions are Gaussian and bi-square, with the bandwidth usually defined through (Euclidean) distance or the number of nearest areas (Mollalo et al., 2020). It should be noted that different bandwidth types would lead to different types of the community, where local weighting happens. R-squared is a measure of goodness of fit. Its value varies from 0.0 to 1.0, and higher values are more preferable. It could be interpreted as the fraction of the variance of the dependent variable that is accounted for by the regression model (Wheeler & Paez, 2010). We interpreted the variables with \( R^2 \) value above 0.5 as being more important in this study.

### 3. Results

#### 3.1. Spatial distribution of COVID-19 in Tehran

We used ArcGIS software and classified the cumulative number of COVID-19 confirmed patients in the following categories: 0; 1–15; 16–25; 26–40; 41–60; and \( N > 60 \) patients per neighborhood (Figs. 3 and 1 of the supplementary appendix). As can be observed in Fig. 3, between 20 February and 20 March, the number of neighborhoods with the cumulative number of COVID-19 cases being in the range of \( 1-15 \), 16-25, 26–40, 40–60, and \( \geq 60 \) persons were 128(36 %), 96(27 %), 89(25 %), 31(9 %), and 10(2 %), respectively. Therefore, there were 169 neighborhoods hit harder by the disease by that date in Tehran (Number of patients > mean), which accounted for 47.6 % of the total number of the neighborhoods and indicated that the pandemic had affected many neighborhoods. On closer inspection based on former categories, the number of neighborhoods from March 20 to April 20 were 90 (25 %), 119 (34 %), 88 (24 %), 46 (13 %), and 14 (4 %), respectively. The classification indicates that, between 20 April and 20 May, 115 neighborhoods (32 %) had less than 15 patients, 124 neighborhoods (35 %) had 16–25 patients, 85 neighborhoods (24 %) had 26–40 patients, 26 neighborhoods (7 %) had 41–60 patients, and seven neighborhoods (2 %) had more than 60 patients. The results of the analysis of the number of patients, based on the same frequency range classification, between 20 May and 20 June and between 20 June and 20 July show that, respectively, there were 209 (59 %) and 98 (28 %) neighborhoods with less than 15 patients, 109 (30 %) and 107 (30 %) neighborhoods with 15–26 patients, 31 (9 %) and 95 (27 %) neighborhoods with 26–40 patients, 8 (2 %) and 19 (15 %) neighborhoods with 41–60 patients, and 19 (5 %) neighborhoods with more than 60 patients.

Concerning the cumulative number of patients in the neighborhoods, the districts with the largest number of confirmed cases in the first month were districts 2, 4, 5, and 15. As with spatial distribution, the parts with higher numbers of confirmed patients were scattered in several areas, including districts 19 and 20. The cumulative number of confirmed patients was highest in districts 4, 5, and 15, while the lowest numbers of confirmed cases were in districts 19 and 20, showing that there was an increasing difference among districts. Regarding spatial distribution, there was an increasing trend of distribution (of infected individuals) in the neighborhoods. In fact, over time, neighborhoods with lower numbers of patients continued to decrease and were predominantly within districts 1, 19, and 20. The spatial distribution of cases is presented in Fig. 3, as cumulative data for five months. Overall, the highest numbers of incidence were in districts 2, 4, 5 and 15.

In comparison, the number of COVID-19 cases was lower in districts 9 and 21. Moreover, a few cases were reported for district 20 (Fig. 1 of supplementary appendix). The maximum numbers of incidence in the first month were recorded in districts 5 and 14. Also, remarkably lower COVID-19 cases were recorded in the fourth month in district 21 (Fig. 1 of supplementary appendix). The heterogeneous distributions of the cases is in line with the pattern of population density of Tehran’s neighborhoods, as shown in Fig. 2.

#### 3.2. Hotspot analysis of patients’ accumulation in Tehran

The analysis showed that the hotspots were mainly distributed in the northeastern, southeastern, eastern, and central Tehran, while less significant hotspots could be observed in northern parts of the city, as shown in Fig. 4. The figure shows that hotspots were initially mainly located in the eastern part of the city, but gradually moved towards the
central parts. Here, hotspots indicate that in some parts patients have agglomerated in proximity with each other. Higher concentration of hotspots in the eastern and central parts may be due to various factors such as lower socio-economic status of these areas that make compliance with social distancing rules challenging. It may also indicate that there are more face-to-face interactions between residents in these parts, unlike the wealthy parts in the north.

3.3. Place-based influencing factors of the pandemic spread

GWR model was used to compute the spatially varying local $R^2$ and coefficient values for each explanatory place-based variable, emphasizing the individual effect of quantitative place-based factors (Figs. 5–7, Table 1). These figures display the profound impact of the components of demographic and urban form on the overall COVID patterns in Tehran’s neighborhoods and districts. Concerning local $R^2$ values for COVID cases, the highest association between demographic variables and the cases was observed for the variable of population density ($R^2 = 0.88$) and a moderate local $R^2$ was observed for people aged above 65 in neighborhoods ($R^2 = 0.31$) (Fig. 4). Concerning demographic factors, all districts showed positive coefficients with population density but the coefficients were higher for districts 4, 8, 10, 13 and 14 (Fig. 6, Table 1).

As for the association between urban form variables and the COVID-19 cases, we developed a ten-parameter regression model employing the local GWR model in three categories of land use, physical setting, and public transportation facilities (Table 1). The strongest association between the explanatory variables related to physical setting and the number of COVID cases was observed for building density ($R^2 = 0.44$). In contrast, the association between the ratio of green spaces to the total area of the neighborhood and the number of confirmed cases was moderate ($R^2 = 0.36$) (Fig. 5, Table 1).

For the land use category, high local $R^2$ values were obtained for neighborhood centers and chain stores ($R^2 = 0.59$, for both), which is indicative of a strong association. Regarding land uses related to healthcare facilities (i.e., drugstores, medical imaging centers, and hospitals), the strongest association with the number of COVID-19 cases...
was obtained for drugstores in neighborhoods ($R^2 = 0.64$). Besides, the association for radiology centers ($R^2 = 0.22$) and hospitals was moderate ($R^2 = 0.24$) (Fig. 6, Table 1).

Among the variables related to public transportation facilities, the highest local $R^2$ was obtained for bus stops, indicating that they are the riskiest places in the neighborhoods ($R^2 = 0.46$). In contrast, the least association was observed between gas stations and the number of confirmed cases ($R^2 = 0.16$). Also, subway stations turned out to have a moderate association ($R^2 = 0.22$) (Fig. 7, Table 1). Overall, among these variables, the most substantial relationship with the number of confirmed cases belongs to frequently-visited locations (drugstores, neighborhood centers, and chain stores). The neighborhood centers are frequently used by residents to meet their daily needs and to socialize, and neighborhood chain stores that are operated by the local municipalities at affordable prices are also often bustling. Also, drugstores that fulfill the healthcare needs of the residents are associated with a higher number of patients that visit on a daily basis and have the highest $R^2$ value.

### 4. Discussion

In this study, we analyzed the spatio-temporal patterns and the place-based influential factors of the COVID-19 pandemic at the level of neighborhoods in Tehran. The results can provide insights into interventions aimed at improving public health policies. We examined the spatial distribution of COVID-19 cases to understand how quantitative place-based factors at the district and neighborhood levels can help identify hotspots, clusters, and potential driving factors. It was found that the distribution of the confirmed cases was heterogeneous across the districts and neighborhoods. This signifies the complex nature of the transmission and spread patterns of the pandemic. In fact, such patterns could be affected by various physical, demographic, cultural, behavioral, and socio-economic factors that may have different impacts across different neighborhoods (Pishgar et al., 2020; Samnigrahni et al., 2020; Sharifi & Khavarian-Garmsir, 2020). There is now ample evidence suggesting that human-to-human interactions during the asymptomatic incubation period of COVID-19 is the main source of virus transmission and spread (Chan et al., 2020; Zhou et al., 2020). Therefore, place-based attributes of urban neighborhoods that affect the extent of human interactions and mobility in a city can influence the transmission and spread patterns of the pandemic (Li et al., 2020). According to the findings of this study, the neighborhood commercial spaces, medical service capacity, and transportation infrastructure in the level of neighborhoods could increase the resident mobility in urban
neighborhoods. In fact, in the absence of appropriate measures such as mobility restrictions and social distancing, high concentration of these facilities may increase the risk of exposure to the virus and cause problems for containing the pandemic. Accordingly, it is critical to take actions towards minimizing congestion in such locations. One possible intervention could be increasing neighborhood self-sufficiency to minimize movement between neighborhoods during adverse events such as pandemics. This could be achieved by following new models such as 15-minute neighborhoods that have recently received considerable attention in planning literature. Such neighborhood-oriented developments can also provide other co-benefits. For instance, they can contribute to social sustainability, support climate change mitigation efforts by reducing travel demand, and foster active travel modes that contribute to physical and mental health of the residents (Sharifi, 2019).

This study also showed that GIS-based spatial modelling facilitates mapping and, potentially, predicting pandemic hotspots and areas with higher risk of virus transmission. According to Fig. 3, in the first five months of the outbreak, the incidence rates were higher in the northern parts of Tehran. Despite the higher socio-economic status of the neighborhoods in districts 2, 4, and 5 (Municipality, 2020), the number of cases were higher in these neighborhoods. This finding contradicts with what is reported for some other parts of the world such as India or the United States, where low-income neighborhoods have been hit harder by the pandemic (Sharifi & Khavarian-Garmsir, 2020). This could possibly be due to the higher availability and accessibility to parks and public open spaces in these districts (Municipality, 2020) or the relatively higher concentration of educational and recreational activities that leads to higher generation and distribution of trips (Municipality, 2020). It should be mentioned that, unlike many other countries that were hit hard by the pandemic, Iran did not implement total lockdown. Therefore, relatively higher use of public and open spaces in these parts

---

Table 1
Summary statistics of the GWR analysis.

| Dimensions | variables                      | P value | Arc  | $R^2$ | $R^2$ Adjusted |
|------------|--------------------------------|---------|------|-------|---------------|
| Place      | Population-over 65             | 0.00    | 3965.12 | 0.31  | 0.21          |
|           | Population Density             | 0.00    | 3468.19 | 0.88  | 0.84          |
|           | Building Density               | 0.00    | 4012.57 | 0.44  | 0.30          |
|           | Green spaces Area Ratio        | 0.00    | 5488.78 | 0.36  | 0.21          |
|           | Neighborhood Shopping and Retail Centers | 0.00 | 5250.02 | 0.59  | 0.55          |
|           | Neighborhood Chain Store       | 0.00    | 4199.96 | 0.59  | 0.53          |
|           | Drugstore                      | 0.00    | 4177.75 | 0.64  | 0.57          |
|           | Radiology centers              | 0.05    | 1354.69 | 0.22  | 0.13          |
|           | Hospitals                      | 0.00    | 1845.52 | 0.58  | 0.24          |
|           | Bus Station                    | 0.00    | 5046.78 | 0.46  | 0.34          |
|           | Metro Station                  | 0.05    | 1354.68 | 0.22  | 0.13          |
|           | Gas stations                   | 0.05    | 2212.16 | 0.16  | 0.14          |

Fig. 7. Spatial pattern of local $R^2$ showing the effects of public transportation facilities in describing COVID-19 incidence rates using the GWR model.
of the city may have contributed to the relatively higher spread of the pandemic, during the first months, in the northern districts. The fact that affluent neighborhoods have also not been immune to the impacts of the pandemic confirms arguments in the literature that socio-economic inequalities may make it more difficult to control the spread of the pandemic, thereby putting the whole city at risk (Sharifi, 2021). Therefore, actions to reduce socio-economic inequalities should be prioritized in the post-COVID era.

Results of the hotspot analysis showed that eastern (4, 8, 13, 14 and 15) as well as central districts (10, 11, 16) are hotspots of the pandemic. In the first five months of the pandemic, the COVID-19 hotspots in all neighborhoods and districts were strongly associated with population density (Figs. 2 and 5). This confirms our first hypothesis that person-to-person transmission of the disease could increase due to population density and the associated increase in the possibility of personal contacts. In addition, in eastern and central districts of Tehran, social ties are stronger and work-related activities associated with economic and commercial sectors are remarkably high (Municipality, 2012, 2020). These may have also contributed to dominance of hotspots in these parts of the city. Overall, the findings showed that higher population density may lead to more cases in the unit of area at the neighborhood scale. In fact, results showed that density has the strongest relationship with the probability of infection. It should, however, be mentioned that better design and enforcement of policies and regulations to maintain social distancing may have mitigated the negative impacts of high-density. However, as mentioned earlier, no strict social distancing policies were implemented in Tehran during the study period. Future research should model potential impacts of such policies on mitigating the negative impacts of density. Also, since work-related activities could be responsible for the dominance of hotspots in the central and eastern districts, well-designed policies to reduce such activities should be implemented during pandemics. Such policies should, however, not negatively impact livelihood capacities of residents, especially in low-income neighborhoods.

The other factor, the population aged above 65 had a moderate association with the infection rate. Previous studies have also confirmed the effect of demographic factors such as the ratio of elderly population on the incidence rate of COVID-19 (Ren et al., 2020; Sharifi & Khavarian-Garmsir, 2020; Xie et al., 2020). Lakhani (2020) found that the people aged above 65 that suffer from other chronic diseases and have limited ability to independently do their household works are particularly vulnerable to COVID-19. Also, Lippi, Mattiuzzi, Sanchis-Gomar and Henry (2020) argued that factors such as being male, aged above 60, and having pre-existing health conditions (e.g., chronic diseases such as diabetes, hypertension, chronic respiratory diseases, cancer, and cardiovascular disorders) strongly determine the rates of COVID-19 infection and death. Similar findings have also been reported in a study in Oman (Mansour et al., 2021).

Among the urban form factors (i.e., physical setting, land use, and public transportation facilities), the strongest relationship with disease rate in neighborhoods belonged to land use. Drugstores had the strongest association with the rate of COVID-19 infection, followed by local shopping and retail centers. Over time, the latter have evolved into retail focal points that have contributed to forming many social relationships. Besides, there are low-price and crowded chain stores in the low-income eastern, central, and southern neighborhoods that can play a significant role in the spread of the disease (Municipality, 2012, 2020). The signifi cance of land use and number of commercial centers (the number of facilities and food services) on the spread of the pandemic has also been confirmed in a study conducted in Huangzhou, China (Li et al., 2020).

Overall, the positive relationship between land use factors and the number of COVID-19 cases shows that these factors could be essential determinants of the overall casualties in neighborhoods. These findings signify the need for better attention to hygiene measures and stricter implementation of social distancing policies in potentially crowded locations such as drugstores, neighborhood shopping and retail centers, and chain stores. In addition, it may be required to take further actions to reduce congestion in such locations during the pandemic. For instance, promoting online shopping can be prioritized in this regard.

Among the physical features of the urban form, the positive relationship between building density and infection rate can be highlighted. This is in line with our earlier discussions about population density and shows that larger numbers of buildings per hectare may increase the odds of transmission by increasing population density. Higher building density implies smaller housing size, and this may further contribute to the disease outbreak by making social distancing challenging. This is particularly the case in the hotspot districts of this study (i.e., east and south), where the mean household size is larger (Municipality, 2020). This results in relatively more congested indoor spaces and densely populated neighborhoods that may increase the infection rate. This is in line with evidence reported in other contexts such as China and India, demonstrating that the population/household density is a risk factor contributing to the transmission of the disease (Ren et al., 2020).

As with the public transportation facilities, bus stops in Tehran have the highest effect on the spread of the disease. Subway stations and gas stations have a moderate relationship with infection rates in neighborhoods. Given the low-income levels of residents of hotspot areas, they are more reliant on the bus transit services. Besides, the bus system has a better geographic coverage and is generally considered as a popular means of local transportation in Tehran. However, although it provides better accessibility and affordability, there are limitations in terms of public health protocols. Evidence from Huangzhou, China also confirms that factor related to transportation networks (the number of bus stops and the length of the road network) have strong effects on the spread patterns of the pandemic (Li et al., 2020). Given the significance of public transportation system for reducing socio-economic and environmental externalities of the transportation sector, it is critical to take necessary actions to overcome potential health risks of public transportation systems.

Although we did not examine the effects of all socio-economic factors in this study, research shows that there are inherent connections between socio-demographic profiles and the number of COVID-19 cases and deaths reported in different parts of the world (Sannigrahi et al., 2020; Sharifi & Khavarian-Garmsir, 2020). The fact that disease hotspots in our study are in districts with lower socio-economic status indicates that such connections may also exist in Tehran. Once again, this highlights the need for taking actions towards reducing socio-economic inequalities in the city. In addition to demographic factors, results of our study are to a large extent aligned with what is reported in the literature. For instance, several factors related to urban form, including distribution patterns of major land use categories such as retail and healthcare centers and public transportation facilities have influenced the spread patterns of the pandemic in Beijing and Guangzhou (Ren et al., 2020).

Our findings related to the impacts of land use and urban form factors imply that some modifications are needed to mitigate the negative impacts of the current and future pandemics in Tehran. Roles and positions of local shopping centers, which once acted as places for socialization and helped to define the neighborhood center may need to be revisited, at least in the time of pandemics. For instance, better penetration and uptake of smart city infrastructure and online platforms (e.g., for shopping, telework, socialization, etc.) may allow better compliance with social distancing measures, while also meeting daily needs of the residents. There is in fact an increasing hope that smart city solutions and infrastructure can help develop urban spaces that are resilient against contagious diseases and contribute to social distancing (Sharifi & Khavarian-Garmsir, 2020).

Overall, results of this study show that place-based attributes can play an essential role in the transmission and spread of infectious diseases in cities. This highlights the significance of careful attention to place-based attributes to develop sustainable and inclusive plans for dealing with health risks, and thereby, contribute to the achievement of Goal 3 (Good Health and Well-being) and Goal 11 (Sustainable Cities...
and Communities) of the Sustainable Development Goals (SDGs). Enhanced knowledge of the role and effects of place-based attributes may allow planners and policy makers to develop measures towards (re-)developing cities in a way that public health concerns could be addressed. This may, however, be time- and resource-intensive, particularly in the already-built cities. Enhanced awareness of risk hotspots may also influence risk perception and behavior of residents and lead to better compliance with social distancing measures (Li et al., 2020).

5. Conclusion

This study examined 43,000 confirmed COVID-19 patients over five months in the neighborhoods of Tehran. We made use of spatial analysis and geographically weighted regression (GWR) methods to analyze the spatio-temporal characteristics and influencing factors of COVID-19 transmission in the city. It is concluded that population density and retail land use in community centers, as well as factors related to health facilities and public transportation such as bus stops have had more effects on the spread of COVID-19 in Tehran. Based on this, there is a need to revisit the way these facilities are distributed in the city and function during pandemics. In fact, some levels of spatial redistribution of urban facilities and population may be needed. This, however, would be challenging and resource-intensive in an already-built city like Tehran. More research is needed to examine possibilities of such interventions. We hope that this effort to identify potential hotspots and risk factors will also inform planners and city authorities of the priority intervention areas and will lead to effective response actions towards better pandemic resilience. Based on the results, it is particularly essential to revisit the role of local shopping centers, reduce potential health risks of the public transportation system, and take initiatives towards reducing intra-city socio-economic inequalities.

While this study has improved our understanding of the spatio-temporal patterns of the pandemic in Tehran and the factors that have contributed to its spread, there are some limitations that should be mentioned to be further addressed in future research. First, not all potentially influential variables have been considered in this study due to limited availability of related place-based and socio-economic data at the neighborhood level in Tehran. A more detailed analysis of socio-economic and also environmental (e.g., pollution, temperature, wind speed, etc.) indicators may also allow us to better understand transmission patterns and identify potential COVID-19 risk zones, and should, therefore, be conducted in the future. Related to data availability, it should also be mentioned that access to data in Iran is challenging. In this study, we could only access data for the first five months of the pandemic. For future research we intend to obtain data for a longer period and also for other Iranian cities. That would allow us to better investigate temporal patterns related to the pandemic. Second, we need to utilize further spatial models (e.g., multilevel geographically weighted regression (MGWR)) to examine the spatial heterogeneities of responses of the infection risks to environmental, socio-economic, and physical variables within the spatial structure of cities, by which the influence of socio-economic conditions on the risk of infection could also be better determined in the risk zones. Third, this study is only based on data from the first five months of the pandemic in Tehran. As the actions and responses of people and local authorities may change over time, follow-up studies that cover longer time periods are needed to obtain more detailed understandings of the spatio-temporal patterns of the pandemic and factors that contribute to its spread. Fourth, although we did not consider the precedent conditions as explanatory variables, they need to be included in future studies such as cardiovascular conditions and diabetes as potential risk factors for the disease. Despite these limitations, we believe that the findings of this preliminary study can be useful for planners and policy makers in their efforts to enhance city preparedness and response and policy development towards controlling the pandemic as well as strategies to create more healthy and sustainable urban neighborhoods.

At the end, we suggest that more studies using the GIS-based spatial modelling approach of this research are needed to better understand the patterns and identify influential demographic and built form factors contributing to the spread of the pandemic. It is, particularly, essential to pay attention to the neighborhood scale as it allows gaining more granular knowledge of spread patterns and influential factors. Despite this, neighborhood-based research is limited particularly in developing countries and the Middle East (Mansour et al., 2021). Spatial modeling not only contributes to discovering where and why clusters and hot spots are located but may also facilitate gaining better knowledge of spatial variations in the incidence and transmission of the disease. This can allow planners and policy makers to implement targeted intervention plans towards enhancing pandemic resilience. As more pandemic-related data will become available in the coming months, we believe that spatial modeling using GWR, used in this study, can be replicated elsewhere in Iran and the Middle East or in other developing countries to better understand the patterns and dynamics of the COVID-19 spread and the associated risk factors, and develop plans and strategies to better respond to the current and future pandemics.

Ethics approval and consent to participate

This article reports a study at Shahid Beheshti University of Medical Sciences with ethics code IR.SBU.RETECH.REC.1399.070. The study did not involve human participants and animals; therefore, no informed consent was needed.

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.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.scs.2021.103034.

References

Alemdar, K. D., Kaya, O., & Çodur, M. Y. (2020). A GIS and microsimulation-based MCDA approach for evaluation of pedestrian crossings. Accident Analysis & Prevention, 148, Article 105771.

Almagro, M., & Orane-Hutchinson, A. (2020). The determinants of the differential exposure to COVID-19 in New York City and their evolution over time. Covid Economics: Vetted and Real-Time Papers, 13.

Barjani, G. J. (2020). Demographic determinants of testing incidence and COVID-19 infections in New York City neighborhoods.

Brunson, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical Analysis, 28(4), 281–298.

Chan, J. F.-W., Yuan, S., Kok, K.-H., To, K. K.-W., Chu, H., Yang, J., et al. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. The Lancet, 395(10223), 514–523.

Fotheringham, A. S., & Ozan, T. M. (2016). Geographically weighted regression and multicollinearity: Dispelling the myth. Journal of Geographical Systems, 18(4), 303–329.

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billi, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. Science of the Total Environment, 739, Article 140033. https://doi.org/10.1016/j.scitotenv.2020.140033

Ghazizadeh, M., Rafieian, M., & Dadaeghoor, H. (2020). Exploring the socio-spatial patterns of diversity and its influencing factors at a metropolitan scale. Journal of Urbanism: International Research on Placemaking and Urban Sustainability, 13(3), 325–356.

Gholamzadeh, S., Mohammadi, G., Looa, M. A., Akrami, F., Zarean, E., Valaei, R., et al. (2020). Baseline Characteristics and Associated Factors of Mortality in COVID-19 Patients; An Analysis of 16000 Cases in Tehran, Iran. Archives of Academic Emergency Medicine, 8(1), e70-e70.

Hao, Y., Yang, L., Jin, K., Li, J., Feng, S., Chen, W., et al. (2021). Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. Science of the Total Environment, 761, Article 144257.

Hu, B., Gong, J., Zhou, J., Sun, J., Yang, L., Xia, Y., et al. (2013). Spatial-temporal characteristics of epidemic spread in-out flow—Using SARS epidemic in Beijing as a case study. Science China Earth Sciences, 56(8), 1380–1397.
