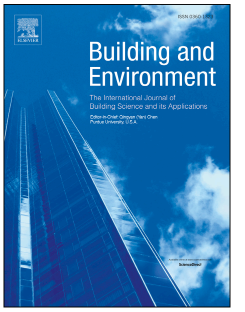
Journal Pre-proof



Relationships between building attributes and COVID-19 infection in London

Huan Tong, Mingxiao Li, Jian Kang

PII: S0360-1323(22)00811-3

DOI: https://doi.org/10.1016/j.buildenv.2022.109581

Reference: BAE 109581

To appear in: *Building and Environment*

Received Date: 5 June 2022

Revised Date: 25 August 2022

Accepted Date: 2 September 2022

Please cite this article as: Tong H, Li M, Kang J, Relationships between building attributes and COVID-19 infection in London, *Building and Environment* (2022), doi: https://doi.org/10.1016/j.buildenv.2022.109581.

Relationships between building attributes and COVID-19 infection in London

Huan Tong[1*], Mingxiao Li[2*], Jian Kang[3*]

[1] School of Architecture, Harbin Institute of Technology (Shenzhen), Shenzhen, China

[2] School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China

[3] Institute for Environmental Design and Engineering, The Bartlett, University College London, London, United Kingdom

*Corresponding Author: Email: huan.tong.18@ucl.ac.uk (Huan Tong), limx@lreis.ac.cn (Mingxiao Li), j.kang@ucl.ac.uk (Jian Kang)

Highlights:

- The uneven spatial distribution ranged from 1837.88 to 4391.79 cases per 10,000 people
- COVID-19 infection rates were lower in higher building density areas, unexpectedly
- Percentage of residents in flats contributed the most to infection rate, negatively

Abstract:

In the UK, all domestic COVID-19 restrictions have been removed since they were introduced in March 2020. After illustrating the spatial-temporal variations in COVID-19 infection rates across London, this study then particularly aimed to examine the relationships of COVID-19 infection rates with building attributes, including building density, type, age, and use, since previous studies have shown that the built environment plays an important role in public health. Multisource data from national health services and the London Geomni map were processed with GIS techniques and statistically analysed. From March 2020 to April 2022, the infection rate of COVID-19 in London was 3159.28 cases per 10,000 people. The spatial distribution across London was uneven, with a range from 1837.88 to 4391.79 per 10,000 people. During the whole COVID-19 control period, it was revealed that building attributes played a significant role in COVID-19 infection. It was noted that higher building density areas had lower COVID-19 infection rates in London. Moreover, a higher percentage of historic or flat buildings tended to lead to a decrease in infection rates. The percentage of residential buildings had a positive relationship with the infection rate. Variations in the infection rate were more sensitive to building type; in particular, the percentage of residents living in flats contributed the most to variations in COVID-19 infection rates, with a value of 2.5%. This study is expected to provide support for policy and practice towards pandemic-resilient architectural design.

1. Introduction

The built environment and population health are intrinsically interlinked in different aspects (e.g., Aletta et al., 2020; Alirol et al., 2011; Matthew & McDonald, 2006; Tong & Kang, 2021; Wang et al., 2022). Historically, cities and buildings have been systematically transformed in response to health threats and other kinds of sanitation issues. Epidemics such as the bubonic plague in the 18th century, cholera in the 19th century, and Spanish flu in the 20th century contributed to sanitary innovations and inspired the value of built environment configurations as important mitigation and prevention strategies (Megahed & Ghoneim, 2020). With rapid urbanisation, current estimates predict that by 2050, two-thirds of the world's population will be living in urban areas (United Nations, 2014). As this occurs, the characteristics of the built environment will play a more important role in promoting public health.

A number of studies have examined the relationship between the built environment and public health from the perspective of infectious disease. For instance, at the urban level, Yashima and Sasaki (2014) indicated that the spread of pandemics is related to the local population size and commuting network structure. Urban green space has positive effects on human health promotion and disease prevention (e.g., Lai et al., 2013; Maas et al., 2006). Moreover, by examining the emergence of past epidemics, Wang et al. (2011) indicated that a city with a multicentric pattern (Shenzhen, China) had fewer cases of severe acute respiratory syndrome (SARS) infection than Hong Kong, which is laid out in a monocentric pattern. Xiao et al. (2014) found that the distribution of influenza H1N1 cases was related to population density and the presence of nearest public places. The H1N1 pandemic was also strongly correlated with urban transportation (Tang et al., 2010). In addition to urban morphology, building attributes, such as ventilation, sanitation, and drainage systems, have been shown to have an impact on virus transmission in high-rise dwellings (e.g., Gao et al., 2008; Lin et al., 2010; Mao & Gao, 2015). Outbreaks of infectious diseases have been inevitable throughout human history. Previous studies have illustrated the importance of urban resistance planning and design in mitigating the impact of disease on population health.

Coronavirus disease 2019 (COVID-19) was first identified in December 2019 and quickly started to affect many regions of the world in the following months (Brown & Horton, 2020). The ongoing COVID-19 pandemic is a global threat to public health, with 519.11 million cases of COVID-19 diagnosed and 6.27 million deaths worldwide as of 16 May 2022 (World Health Organization, 2022). Key characteristics of the urban built environment, such as urban morphologies and building attributes, have been documented to have an impact on COVID-19 infection. For instance, at the urban level, city size is a key factor influencing the transmission of viral disease in US cities, with COVID-19 spreading faster on average in larger cities (Stier et al., 2020). AbouKorin et al. (2021) found that radial and grid cities were associated with higher rates of COVID-19 infection than linear cities. Moreover, green space was found to reduce the impact of COVID-19 transmission and lower COVID-19 mortality rates (Frank & Wali, 2021; Sallis et al., 2022). In addition, there is evidence that socioeconomic factors, including income, unemployment rate, education level, health status, race, and other characteristics, are contributors to COVID-19 infection (e.g., Almagro

& Orane-Hutchinson, 2020; Drefahl et al., 2020; Lei et al., 2020; Khunti et al., 2020; Pareek et al., 2020). Raifman and Raifman (2020) indicated that income was strongly related to virus exposure in the United States. Mollalo et al. (2020) also suggested a significant correlation between income and COVID-19 incidence rates. Meanwhile, they concluded that COVID-19 infection was related to outdoor environmental factors, including road density, particulate matter 2.5, air quality index, temperature, and precipitation. At the building level, Kwok et al. (2021) found that building density has a substantial effect on COVID-19 infection by examining COVID-19 in Hong Kong. They found that building height can lead to an increased risk of COVID-19 infection. The indoor environment, including occupants, ventilation, and indoor air quality, was also related to the spread of COVID-19 (e.g., Dietz, 2020; Eykelbosh, 2020). However, research on the role of building attributes, such as the building type and use, in COVID-19 infection is still lacking.

In England, the virus began circulating in early 2020. To mitigate its impact, the UK government passed the Health Protection (Coronavirus, Restrictions) (England) Regulations 2020, which were implemented at 1:00 pm on 26 March 2020 (Public Health England, 2020). Subsequently, with the new variants of COVID-19, lockdown measures have been changed accordingly. On 24 February 2022, all domestic COVID-19 restrictions were lifted in England under the government's announced "living with COVID" plan. This offers a good opportunity to analyse COVID-19 infection during the whole control period, although there is still a large population infected with COVID-19.

Therefore, by first examining the spatial-temporal distribution of COVID-19 infection from 14 March 2020 to 22 April 2022, this study then particularly aimed to investigate the relationships between COVID-19 infection and building attributes, including building density, type, age, and use. To address this issue, building geometry data and infection case data from the governmental open data platform were processed with geographic information system (GIS) techniques. Multivariate linear regression was used to model COVID-19 infection rates and building attributes, simultaneously adjusting for socioeconomic factors. It is expected that the results can inform architecture design and urban planning to build a healthy and resilient city able to withstand future pandemics.

2. Methods

2.1. Case study site

Greater London has a population of approximately 8.9 million and a population density of 64.16 people/hectare (Figure 1). There are 982 Middle Layer Super Output Areas (MSOAs), a geographical hierarchy designed to improve the reporting of small area statistics in England and Wales (NHS Data Model and Dictionary, 2022). There are several reasons that make London well suited for a case study. First, the COVID-19 infection, socioeconomic factor, and building attribute datasets have the same data collection methods and have available information from across London. There is a wide variation in building attributes, and infection rates vary considerably across London. Moreover, as mentioned above, lockdown measures have been in place in London since March 2020, and all COVID-19 restrictions have

Journal Pre-proof

now ended. This marks a new phase in the COVID-19 pandemic. It is therefore an opportune time to investigate the role of the built environment in COVID-19 infection. Moreover, public health policy responses to the COVID-19 pandemic are consistent across London. Therefore, this study focused on London and selected 981 MSOAs for analysis (excluding the City of London because data were not available).

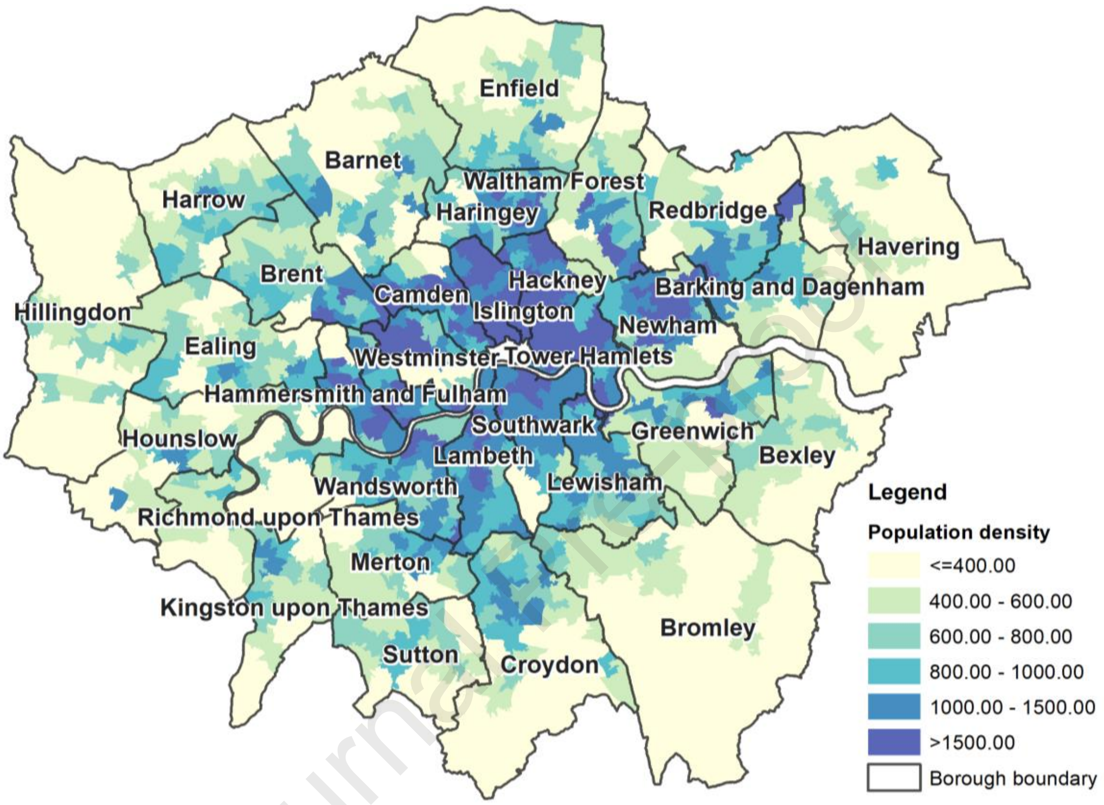


Figure 1 The distribution of population density (people per square kilometre) in London.

2.2. Data sources and indicators

This study involved COVID-19 infection, building attribute, and socioeconomic factor datasets in London. COVID-19 infection data were obtained from the UK Coronavirus Dashboard, which was developed by the UK Health Security Agency (2022). The dashboard is a timely and authoritative summary of key information about the COVID-19 pandemic and includes levels of infection cases, testing, deaths, and vaccination data. The dashboard supports researchers in reusing data by accessing results in machine-readable files and via an application programming interface (UK Health Security Agency, 2022). The number of COVID-19 infection cases was obtained for the rolling 7-day period from 14 March 2020 to 22 April 2022, which covers the start and end of COVID-19 restrictions in London. COVID-19 infection data were available for different geographical areas, such as nations, English regions, local authorities, and MSOAs. Among these, infection cases in MSOAs were the most readily available at the local level. The infection rate was calculated by the number of people infected with COVID-19 per 10,000 people.

Based on previous studies, COVID-19 infection rates are potentially related to several built environment indicators, including dwelling type, physical morphology (density and height), and land use. Given the availability of data, the 19 building indicators obtained were categorised into building density, type, age, and use, which are important building attributes. Building density, type, and age were obtained from the UKBuildings dataset in 2021, a national database of building features developed and maintained by Geomni that provides detailed information about individual buildings across the UK. Building use was recategorised according to building use values from the original UKBuildings dataset (Table S1 in Supplementary File). Defence, storage, utility, and unclassified buildings were not included in the analysis due to limited data and missing information. UKBuildings is a spatial dataset that records the location and footprints of buildings with several attributes that describe building features. Then, the datasets were processed in ArcGIS 10.4 to calculate the values of the indicators of buildings at the MSOA level with the help of the spatial analysis, attribute link, and spatial statistics modules. In addition, building types were sourced from the UK Census and summarised to the MSOA level. The building indicators are described in Table 1, and the corresponding calculation method is also illustrated.

Table 1 Indicators for building attributes and descriptions.

Category	Indicators	Descriptions
Building density	Floor area ratio	The ratio of the building's total floor area to the area of the MSOA
	Building base density	The base area of the building to the area of the MSOA
Building type	Detached house	The percentage of residents living in detached houses in a particular MSOA
	Terraced house	The percentage of residents living in terraced houses in a particular MSOA
	Flat	The percentage of residents living in flats in a particular MSOA
Building age	Historic building	The percentage of historic buildings
	Interwar building	The percentage of interwar buildings
	Postwar building	The percentage of postwar buildings
	Sixties-seventies era building	The percentage of sixties-seventies era buildings
	Modern building	The percentage of modern buildings
Building use	Community building	The percentage of community buildings
	Commercial building	The percentage of commercial buildings
	Industry building	The percentage of industrial buildings
	Office building	The percentage of office buildings
	Recreation and leisure building	The percentage of recreation and leisure buildings
	Retail building	The percentage of retail buildings
	Residential building	The percentage of residential buildings
	Transport building	The percentage of transport buildings
	Agricultural building	The percentage of agricultural buildings

COVID-19 infection is also affected by socioeconomic factors, such as demographic, environmental, social, and transportation factors, in urban contexts (e.g., AbouKorin et al., 2021; Baena-Díez et al., 2020; Sharifi & Khavarian-Garmsir, 2020). Therefore, socioeconomic factors were included as control variables in this study. Based on previous studies, population density, median age, income, ethnicity, employment, students, education, health, crime, local services, living environment, and transportation mode were considered and included. Considering the elimination of multicollinearity between the control variables and the balance of model performance and the number of variables entered, three indicators were finally extracted: population density, deprivation index, and the percentage of people who commuted by public transport. In particular, the deprivation index encompasses a wide range of individual living conditions and generally represents the socioeconomic status of an area (Ministry of Housing, Communities and Local Government, 2019). The socioeconomic factor dataset was obtained from the Office of National Statistics (Ministry of Housing, Communities and Local Government, 2019; Park, 2021). In addition, in London, the vaccination rate was strongly related to the deprivation index. The deprivation index, which has been included in the models as a key control variable, can present the level of vaccination rate across different areas in London. Therefore, the vaccination rate was not entered directly, which was also due to multicollinearity issues.

2.3. Statistical analysis

Multivariate linear regression, which is one of the most widely used techniques in built environment research, was chosen to model COVID-19 infection rates and building attributes (e.g., French et al., 2014; Ma & Dill, 2015; Moghadam et al., 2018). Simultaneously, socioeconomic factors were adopted as control variables. In this study, COVID-19 infection rates, building attributes, and socioeconomic factors were continuous variables, which is the essential assumption of the multivariate linear regression model. From the scatter plots, linear relationships between the dependent variable and each independent variable were generally observed. However, combined with casewise diagnostics, 12 cases were identified as outliers and eliminated as they were distant from other cases. The results for multicollinearity, independent errors, homoscedasticity, and normally distributed residual checks are presented in Section 3.2.2.

In this study, the COVID-19 infection rates were modelled as a function of building attributes and socioeconomic factors by using a multivariate linear regression framework. The statistical model was given as

$$Y_k = \beta_0 + \beta_1 * \text{building}_k + \gamma_j * \sum_{j=1}^{m} s_{jk} \quad (1)$$

where Y_k is the observed COVID-19 infection rate (the number of COVID-19 infection cases per 10,000 people) in MSOA k; $building_k$ is the indicator of building attributes in MSOA k; and s_{jk} is the control variable (m = 3, three socioeconomic factors were retained). Moreover, the multivariate regression analysis was subsequently conducted in Statistical Package for the Social Sciences 28.0 (IBM Corp, 2015). Due to the high correlation between indicators of building attributes, multiple building attributes were tested in separate regression models.

3. Results

3.1. Spatial and temporal distribution characteristics of COVID-19 infection rate

From 14 March 2020 to 16 April 2022, there were 2,822,986 COVID-19 infections, with an infection rate of 3159.28 cases per 10,000 people in London. The spatial distribution of the COVID-19 infection rate in

London is shown in Figure 2. COVID-19 infection rates were not evenly distributed, with an infection rate range (namely, the difference between the lowest and highest) of 2553.91 cases per 10,000 people and a standard deviation of 342.41. The highest COVID-19 infection rate was observed in Acre Lane in Lambeth Borough, with a value of 4391.79 per 10,000 people. In contrast, the lowest rate of infection was observed in Knightsbridge, Belgravia and Hyde Park in Westminster, at 1837.88 per 10,000 people.

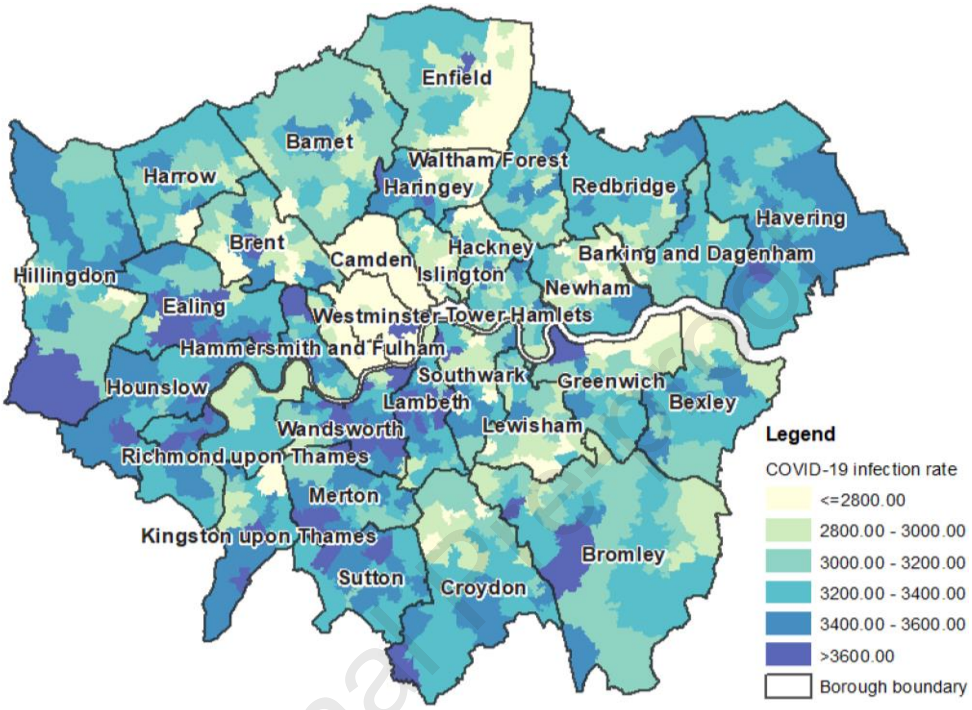


Figure 2. Spatial distribution of COVID-19 infection rates (the number of people infected with COVID-19 per 10,000 people) across London.

In terms of the temporal distribution of COVID-19 infection rates from 2020 to 2022 (Figure 3), the infection rates were not constantly increasing. Instead, two peaks were observed. The first peak occurred in the rolling 7-day period ending on 2 January 2021, when the Alpha variant was identified and swept rapidly across the UK (Grint et al., 2021; Ladhani et al., 2021). The rate was 106.33 cases per 10,000 people. The second peak was observed on 25 December 2021, with an infection rate of 208.02, when Omicron was identified and replaced Delta as the predominant variant (Paton et al., 2022). It can be seen that the rates of COVID-19 infection were increased across all the boroughs in London after the Omicron was identified. Moreover, a less obvious peak occurred in July 2021, when the Delta variant became the dominant variant and swept the UK (Torjesen, 2021). The trends in the infection rates over time were similar across London.

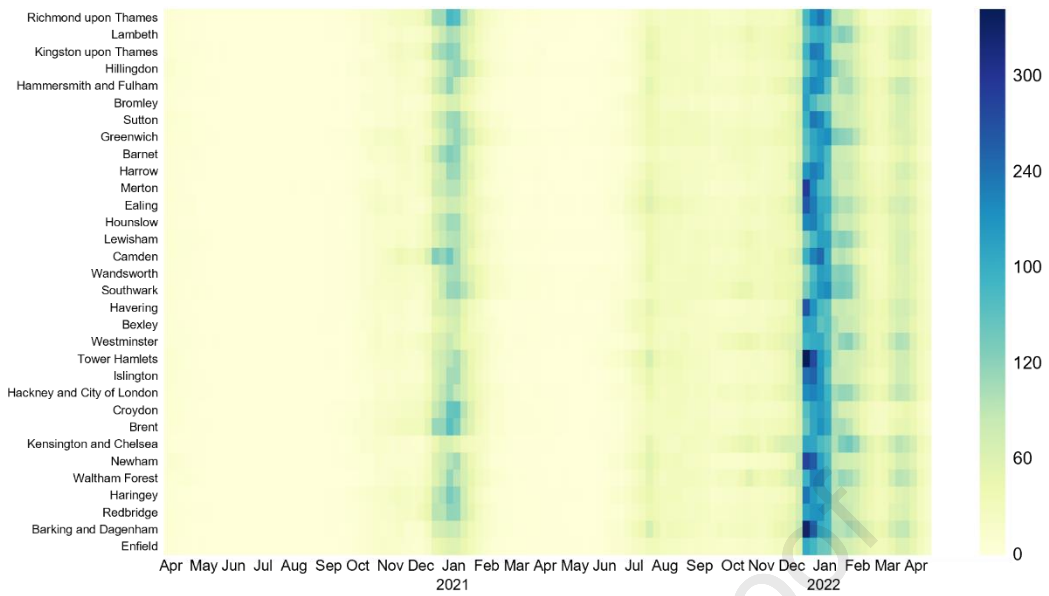


Figure 3. Temporal distribution of weekly COVID-19 infection rates (the number of people infected with COVID-19 per 10,000 people).

3.2. Multivariate linear regression analysis results

When only socioeconomic factors (i.e., control variables including population density, deprivation index, and the percentage of people commuting by public transport) were entered into the multivariate linear regression, the model was significant (F-statistic value less than 0.001), meaning that the independent variables (i.e., control variables) influenced the dependent variable (COVID-19 infection rates). Accordingly, an R square of 0.199 indicated that the control variables could explain 19.9% of the variance in COVID-19 infection rates. Moreover, each building attribute indicator was added into the regression model one-by-one, and the results are presented below. Overall, during the whole process of the COVID-19 control period, building attributes explained the variation in COVID-19 infection rates across London to a different extent.

3.2.1. Effect of building attributes

The results of the statistical model for building density and COVID-19 infection rates are presented in Table 2, which shows the estimated association between the COVID-19 infection rate and the two key building density indicators on a regression coefficient scale. After adjusting for multiple socioeconomic factors, the COVID-19 infection rate was negatively related to the floor area ratio, with the COVID-19 infection rate tending to decrease by 134.59 cases per 10,000 people as the floor area ratio increased by one unit. The floor area ratio and socioeconomic variables explained 21.5% of the variations in London COVID-19 infection rates. The floor area ratio explained an additional 1.6% of the variations. No significant relationship was found between building base density and the COVID-19 infection rate.

Table 2. Regression coefficients and 95% confidence intervals for the COVID-19 infection rate associated with building density.

Indicators	Regression coefficients	95% confidence intervals for coefficients		Significance level	Cumulative R square	Additional R square
		Lower bound	Upper bound			
Floor area ratio	-134.59	-192.96	-76.22	<0.001**	0.215	0.016
Building base density	1.55	-4.72	7.82	0.629	0.199	0.000

* Regression is significant at the 0.05 level.
** Regression is significant at the 0.01 level (Following notes are same).

In terms of building type, the regression analysis results of the COVID-19 infection rate are shown in Table 3. After adjusting for socioeconomic factors, all building type indicators were significantly related to COVID-19. Whole houses (detached houses and terraced houses) had positive relationships with COVID-19 infection rates. The percentage of residents living in detached houses was estimated to contribute to 1.81 additional cases of COVID-19 per 10,000 people at the 0.019 significance level, whereas the percentage of residents in terraced houses contributed to 2.54 additional infection cases per 10,000 people at the 0.001 significance level. In terms of flats, with an increasing percentage of residents living in flats, there was a decrease in the COVID-19 infection rate. An extra one percent increase in residents living in flats was likely to decrease the COVID-19 infection rate by 3.07, which was higher than that for the other types of buildings. The flat variable explained an additional 2.3% of the variations in COVID-19 infection rates.

Table 3. Regression coefficients and 95% confidence intervals for the COVID-19 infection rate associated with building type.

Indicators	Regression coefficients	95% confidence intervals for coefficients		Significance level	Cumulative R square	Additional R square
		Lower bound	Upper bound			
Detached house	1.81	0.30	3.32	0.019*	0.203	0.004
Terraced house	2.54	1.29	3.78	<0.001**	0.212	0.013
Flat	-3.07	-4.20	-1.93	<0.001**	0.222	0.023

In terms of building age, Table 4 shows the association between the age of buildings and COVID-19 infection rates. In the multivariate models, the only variables associated with COVID-19 infection rates that remained significant were historic and interwar buildings. An additional percentage increase for historic buildings was related to 2.75 fewer cases of COVID-19 infection per 10,000 people, explaining 2.1% of the variance in the COVID-19 infection rate. The percentage of interwar buildings was positively related to COVID-19 infection rates, with an increase of one percent of interwar buildings leading to another 1.76 infections per 10,000 people. In terms of the percentage of postwar, sixties-seventies era, and modern buildings, no significant relationship with the COVID-19 infection rate was found.

Table 4. Regression coefficients and 95% confidence intervals for the COVID-19 infection rate associated with building age.

Indicators	Regression coefficients	95% confidence intervals for coefficients	Significance level	Cumulative R square	Additional R square

		Lower bound	Upper bound			
Historic building	-2.75	-3.80	-1.70	<0.001**	0.220	0.021
Interwar building	1.76	0.26	3.25	0.021*	0.203	0.004
Postwar building	2.91	-0.51	6.23	0.095	0.201	0.002
Sixties-seventies era building	-1.32	-3.95	1.31	0.324	0.200	0.001
Modern building	1.07	-1.47	3.61	0.407	0.199	0.000

Table 5 shows the results of each multivariate model for COVID-19 infection rates and building use. The percentages of community, commercial, and industrial buildings was not significantly related to COVID-19 infections. The percentages of office, recreation and leisure, and retail buildings were negatively correlated with COVID-19 infection rates, with each type of building use explaining 0.6% of the variance in the rates after adjustment for socioeconomic factors. Specifically, an increase in the percentages of office and retail buildings was estimated to reduce COVID-19 infection rates by 3.61 and 3.31, respectively. The percentage of recreation and leisure buildings had a slightly greater influence on the magnitude of increase in COVID-19 infection rates than office and retail buildings. Each unit increase in the percentage of recreation and leisure buildings tended to contribute to a 6.44% decrease in infection rates. A positive relationship was observed for residential buildings, whereby COVID-19 infection rates tended to increase by 1.73 per 10,000 people as the percentage of residential buildings increased. The percentage of residential buildings explained an additional 0.9% of the variance in COVID-19 infection rates, with a slightly higher contribution than other building use types. Finally, no significant relationship was observed for transport or agricultural buildings.

Table 5. Regression coefficients and 95% confidence intervals for the COVID-19 infection rate associated with building use.

Indicators	Regression coefficients	95% confidence intervals for coefficients		Significance level	Cumulative R square	Additional R square
		Lower bound	Upper bound			
Community building	-0.98	-3.72	1.76	0.482	0.199	0.000
Commercial building	0.68	-1.54	2.91	0.546	0.199	0.000
Industry building	-1.68	-5.92	-2.57	0.438	0.199	0.000
Office building	-3.61	-6.17	-1.04	0.006**	0.205	0.006
Recreation and leisure building	-6.44	-11.09	-1.80	0.007**	0.205	0.006
Retail building	-3.31	-5.67	-9.40	0.006**	0.205	0.006
Residential building	1.73	0.69	2.76	<0.001**	0.208	0.009
Transport building	-2.23	-7.49	3.03	0.405	0.199	0.000
Agricultural building	2.52	-16.18	21.22	0.792	0.199	0.000

3.2.2. Control variables and assumption checks

As mentioned above, socioeconomic factors were considered control variables, and the model was significant. Moreover, as expected, all socioeconomic variables were relatively consistent across all regression models in terms of the magnitude and significance level of the coefficients. Socioeconomic factors explained 19.9% of the variance in COVID-19 infection rates: population density and deprivation index were negatively related to COVID-19 infection rates at the 0.001 significance level, whereas the percentage of people commuting by public transport had a positive relationship with COVID-19 infection rates at a significance level of 0.001.

When modelling the relationship between COVID-19 infection rates and building attributes, the assumptions of the multivariate linear regression model were also checked, including multicollinearity, independent errors, homoscedasticity, and normally distributed residuals. Multicollinearity, one of the essential hypotheses for multivariate regression analysis, occurs when the independent variables in a regression model are highly correlated. Multicollinearity can lead to modelling problems, such as a reverse sign or wider confidence intervals of the regression coefficients, which could cause misleading interpretations of modelling results (Gregorich et al., 2021). To detect multicollinearity, variance inflation factors (VIFs) were used and analysed to check for any multicollinearity issues in the multivariate regression models. For continuous variables, a VIF greater than 10 indicates that the independent variables are highly correlated (AbouKorin et al., 2021). In this study, the VIFs for all variables in the multivariate regression models were less than 2. Therefore, there was no multicollinearity issue. Moreover, in terms of independent errors, the Durbin-Watson statistic showed that the values of the residuals were slightly positively autocorrelated. However, the Durbin-Watson statistic value (approximately 1.2) fell in the range of 1 to 3, which was acceptable. Values below 1 and above 3 can cause concern and may invalidate the analysis. Furthermore, in terms of homoscedasticity and residual distribution, scatter plots of standardised residuals vs. standardised predicted values showed no obvious signs of funnelling, and the P-P plot for all models in this study showed that residuals were normally distributed (University College London, 2022).

4. Discussion

4.1. The effect of building attributes on COVID-19 infection rates

London faced a profound public health crisis with a rate of 3159.28 COVID-19 infections per 10,000 people. However, the infection rate in London was lower than the average value in England at 3,299.12 per 10,000 people and ranked second to last across regions in England (UK Health Security Agency, 2022). Meanwhile, the distribution of infection rates was spatially and temporally uneven. The tendency over time, as expected, corresponded to the outbreaks of new COVID-19 variants. In this study, building attributes were examined via statistical analysis. In general, several building attribute indicators were related to the COVID-19 infection rate.

High building density was negatively associated with the COVID-19 infection rate. This finding was somewhat counterintuitive since crowded living conditions should accelerate the spread of COVID-19 due to frequent face-to-face interaction. However, existing studies show that evidence for the relationship between density and COVID-19 was contradictory and inconclusive. Pafka (2020) stated that although physical distancing is the most common measure to contain the spread of the virus, this does not mean that higher density areas necessarily have more COVID-19 cases and lower density areas are more resilient to the pandemic. Boterman (2020) did not find a significant relationship between density and the rate of COVID-19 infection in the Netherlands. Similarly, in an investigation of over 900 US metropolitan counties, Hamidi et al. (2020) found that density was not linked to rates of COVID-19 infection. Surprisingly, COVID-19 death rates are significantly lower in high-density counties. The reason for this is difficult to explain. However, as indicated by previous research, in addition to better accessibility to health care facilities, dense areas may be better environments for taming and enforcing strict measures and easier management of social distancing interventions (Hamidi et al., 2020). Moreover, in dense buildings, the coverage of high-speed internet and home delivery services is highly available; hence, residents can conveniently stay at home and avoid unnecessary contact with others (Fang & Wahba, 2020).

Subsequent analysis of building age and type supported this finding for building density and COVID-19 infection: detached/terraced houses or historic buildings, which are generally found in low-density areas, tended to increase the COVID-19 infection rate, whereas flats (typically found in areas exhibiting high density) had a negative relationship with the infection rate. A possible explanation for this result is social factors, such as age and families with students/pupils. Previous studies have found that age impacts the infection rate and that families with students/pupils were also affected by COVID-19 (Ehlert, 2021; Emeruwa et al., 2020; Lei, 2020). However, these indicators were considered as control variables; hence, the results might not be caused by the factors of age and families with students/pupils. Furthermore, household size is also an important socioeconomic factor, and previous studies have indicated a significant association between large households and COVID-19 infection (Ehlert, 2021; Emeruwa et al., 2020). Therefore, to test this, more analyses were conducted. Household size was added to the multivariate linear regression model in an attempt to explain the results. It is found that household size was not significantly related to the COVID-19 infection rate and the R square was almost the same as that in the model without household size with the value of 0.223. Therefore, household size did not explain the reduction effect of the percentage of flats on COVID-19 infection rates in this study. Another possible reason is that this result could be related to the structure of flats in London, a great number of which have exterior stairs and corridors connecting flat entrances, which largely avoid face-to-face interaction. Moreover, these flats have separate and well-constructed natural ventilation systems. Such a layout makes the infection rate in flats not as high as expected. If the flats have individual mechanical ventilation systems, the infection rate would be as low as that of the flats with natural ventilation. However, flats with individual mechanical ventilation systems are rather limited in London. It is expected that these factors can be considered in resistance building design to mitigate the impact of pandemics. Overall, among building attributes, building type explained more of the COVID-19 infection rates; in particular, the percentage of flats contributed the most to variations, with a value of 2.5%.

In terms of building use, in public buildings, such as those for offices, leisure, and retail, the COVID-19 infection rate tended to be lower, whereas residential buildings were likely to have higher infection rates. These relationships may be explained by “stay at home” lockdown measures, which were recommended for residents. Residents were not allowed to leave their homes or go to public buildings, which were closed during the unexpected period under the strictest restrictions (Public Health England, 2020). Hence, public buildings had a relatively low COVID-19 infection rate, whereas residential buildings had a relatively high rate.

4.2. Implications

Building attributes played an important role in the spread of COVID-19. The COVID-19 infection rates varied according to building density, type, age, and use. These findings can inform architectural design from the perspective of pandemic-resilient buildings. For instance, designers can pay more attention to improving the performance of interwar buildings. In addition to general maintenance, the indoor environment and sanitation should also be improved to prevent disease transmission. Moreover, despite the lockdown measures being lifted, the events of the COVID-19 pandemic did change people's daily lives and views on working from home, which is likely to become an increasingly common practice in the future. Combined with the positive relationship between residential buildings and the COVID-19 infection rate, emphasis should be placed on the performance of residential buildings, such as improving ventilation conditions, which has been shown to have an impact on virus transmission (e.g., Bhagat et al., 2020; Li et al., 2021; Xu et al., 2020). A good living environment can also benefit the mental health of residents during the lockdown period. The built environment is important to mitigate the impact of disease on population health and societal development. Architectural design, as a nonpharmaceutical intervention, plays an essential role in preventing pandemics and eliminating virus transmission.

4.3. Limitations and future research

This study suggests a number of limitations and possibilities for future research. The first aspect to consider is related to the COVID-19 dataset. The COVID-19 dataset obtained from UK open data was limited by testing capacity and willingness to test. Although these public data were widely used in a number of previous studies (e.g., Anderson et al., 2020; Ghosh et al., 2020), it would be better to have more information on the actual number of COVID-19 infection cases. Second, drawing from the statistical models, the Durbin-Watson statistic values for the built regression models were approximately 1.2, indicating that independent variables (i.e., COVID-19 infection rates) were slightly positively autocorrelated. In this study, this meant that a spatial correlation (i.e., the neighbourhood effects) existed between buildings for COVID-19 infections. Although Durbin-Watson values in the range of 1-3 are acceptable, it would be interesting to investigate the effects of spatial relationships among buildings in future studies. The indicators for the spatial relationship among buildings, such as the distance between buildings, could be included. Third, UK have experienced different lockdown periods, such as first national lockdown, minimal lockdown restrictions, reimposing restrictions, second national lockdown, and other lockdown periods (UK Parliament, 2021). Even though this study focused on the whole period from March 2020 to April 2022, it is worth investigating the impact of different periods. However, the delayed nature of policy implementation and the complexity of human behaviour in response to policy make it difficult to clearly divide the different time periods. As an attempt, this study conducted an additional analysis for the relatively strict lockdown period (from March 2020 to March 2021) and leaving lockdown period (from March 2021 to February 2022) (Tables S2 and S3 in Supplementary File). It is found that during strict lockdown period, the results were similar to the analysis for the whole period, i.e. the regression coefficients have almost the same sign and similar magnitude. During leaving lockdown period, the results seemed to be somewhat different: some coefficients became not significant and the coefficient values became lower. Therefore, the results might be different, due to division of the time period. Therefore, the precision of the time period division may lead to different results. In future investigations, based on more standard and detailed division of time period criteria, the effects of different lockdown periods could be studied in more depth. Fourth, in this study, the control variables (i.e., socioeconomic factors) were significantly related to the COVID-19 infection rate. From this perspective, it would also be useful to consider other cities where the socioeconomic conditions (e.g., ethnicity) are different. Finally, the variation of COVID-19 is always mutating and more strains may emerge in the future. The transmission characteristics of each strain are different. Therefore, it would be interesting to investigate the impact of different variations. For instance, omicron, one of the most important variations, differs from other variations due to its highly contagious. Omicron is still popular and there are also new strains. In the future, a further study with more focus on Omicron is therefore suggested if the data of Omicron infection rate is available.

5. Conclusions

Based on multisource data, GIS techniques, and statistical analysis, this study is the first to illustrate the spatial-temporal distribution of COVID-19 infection rates, particularly regarding the relationship between COVID-19 infection rates and building attributes. From March 2020 to April 2022, the infection rate of COVID-19 in London was 3159.28 cases per 10,000 people, which was lower than the average in England. The tendency over time corresponded to the outbreaks of new COVID-19 variants.. Moreover, the spatial distribution of infection rates across London was uneven, with a range from 1837.88 to 4391.79 per 10,000 people.

These results revealed that throughout the control period of COVID-19, building attributes played a significant role in COVID-19 infection. In general, a number of building attribute indicators contributed to variations in the COVID-19 infection rate. Areas with higher building density were more likely to have a lower infection rate in London. Meanwhile, the higher percentage of historic or flat buildings tended to lead to a decrease in infection rates. In terms of building use, the rate of COVID-19 infection tended to be lower in public buildings and higher in residential buildings. The variations in COVID-19 infection rates were more sensitive to building type. In particular, the percentage of residents living in flats explained an additional 2.5% of the COVID-19 infection rate variations and contributed the most among all the building attributes.

In addition, as previous studies have indicated, it is expected that the spread of COVID-19 would be related to control variables, i.e., socioeconomic factors. In checking the assumptions of the model, the spatial relationship among buildings (e.g., the distance between buildings and degree of building enclosure) had an effect on COVID-19 infection rates, for which further research could be carried out.

Despite the removal of COVID-19 restrictions, the dramatic events of 2020 did change people's daily lives and raised their awareness of future crises and upcoming pandemics. Working from home is likely to become an increasingly common practice in the future. Buildings, especially low-density residential buildings, will then be an even more crucial living environment in terms of disease prevention and mental health promotion. This study is expected to be useful for policy and practice in pandemic-resilient architectural design.

6. References

AbouKorin, S. A. A., Han, H., & Mahran, M. G. N. (2021). Role of urban planning characteristics in forming pandemic resilient cities–Case study of Covid-19 impacts on European cities within England, Germany and Italy. Cities, 118, 103324.

Aletta, F., Oberman, T., Mitchell, A., Tong, H., & Kang, J. (2020). Assessing the changing urban sound environment during the COVID-19 lockdown period using short-term acoustic measurements. Noise Mapping, 7(1), 123-134.

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

Anderson, R. M., Hollingsworth, T. D., Baggaley, R. F., Maddren, R., & Vegvari, C. (2020). COVID-19 spread in the UK: the end of the beginning?. The Lancet, 396(10251), 587-590.

Baena-Díez, J. M., Barroso, M., Cordeiro-Coelho, S. I., Díaz, J. L., & Grau, M. (2020). Impact of COVID-19 outbreak by income: hitting hardest the most deprived. Journal of Public Health, 42(4), 698-703.

Bhagat, R. K., Wykes, M. D., Dalziel, S. B., & Linden, P. F. (2020). Effects of ventilation on the indoor spread of COVID-19. Journal of Fluid Mechanics, 903.

Boterman, W. R. (2020). Urban‐rural polarisation in times of the corona outbreak? The early demographic and geographic patterns of the SARS‐CoV‐2 epidemic in the Netherlands. Tijdschrift voor economische en sociale geografie, 111(3), 513-529.

Brown, A., & Horton, R. (2020). A planetary health perspective on COVID-19: a call for papers. The Lancet, 395(10230), 1099.

Dietz, L., Horve, P. F., Coil, D. A., Fretz, M., Eisen, J. A., & Van Den Wymelenberg, K. (2020). 2019 novel coronavirus (COVID-19) pandemic: built environment considerations to reduce transmission. Msystems, 5(2), e00245-20.

Drefahl, S., Wallace, M., Mussino, E., Aradhya, S., Kolk, M., Brandén, M., ... & Andersson, G. (2020). Socio-demographic risk factors of COVID-19 deaths in Sweden: A nationwide register study. Stockholm Research Reports in Demography, 23.

Ehlert, A. (2021). The socio-economic determinants of COVID-19: A spatial analysis of German county level data. Socio-Economic Planning Sciences, 78, 101083.

Emeruwa, U. N., Ona, S., Shaman, J. L., Turitz, A., Wright, J. D., Gyamfi-Bannerman, C., & Melamed, A. (2020). Associations between built environment, neighborhood socioeconomic status, and SARS-CoV-2 infection among pregnant women in New York City. Journal of American Medical Association, 324(4), 390-392.

Eykelbosh, A. (2020). Physical barriers for COVID-19 infection prevention and control in commercial settings. National Collaborating Center for Environmental Health: Vancouver, BC, Canada.

Fang, W., & Wahba, S. (2020). Urban density is not an enemy in the coronavirus fight: Evidence from China. Retrieved from World Bank blogs: https://blogs.worldbank.org/sustainablecities/urban-density-not-enemy-coronavirus-fight-evidence-china

Frank, L. D., & Wali, B. (2021). Treating two pandemics for the price of one: Chronic and infectious disease impacts of the built and natural environment. Sustainable Cities and Society, 73, 103089.

French, S., Wood, L., Foster, S. A., Giles-Corti, B., Frank, L., & Learnihan, V. (2014). Sense of community and its association with the neighborhood built environment. Environment and Behavior, 46(6), 677-697.

IBM Corp. (2015). IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.

Gao, N. P., Niu, J. L., Perino, M., & Heiselberg, P. (2008). The airborne transmission of infection between flats in high-rise residential buildings: tracer gas simulation. Building and Environment, 43(11), 1805-1817.

Geomni UK. (2021). UKBuildings [GeoPackage geospatial data] EDINA Geomni Digimap Service, https://digimap.edina.ac.uk

Gregorich, M., Strohmaier, S., Dunkler, D., & Heinze, G. (2021). Regression with highly correlated predictors: variable omission is not the solution. International Journal of Environmental Research and Public Health, 18(8), 4259.

Grint, D. J., Wing, K., Houlihan, C., Gibbs, H. P., Evans, S. J., Williamson, E., ... & Eggo, R. M. (2021).

Severity of SARS-CoV-2 alpha variant (B. 1.1. 7) in England. Clinical Infectious Diseases, ciab754.

Ghosh, A., Nundy, S., Ghosh, S., & Mallick, T. K. (2020). Study of COVID-19 pandemic in London (UK) from urban context. Cities, 106, 102928.

Katella, K. (2022). 5 things to know about the delta variant. https://www.yalemedicine.org/news/5-things-to-know-delta-variant-covid.

Khunti, K., Singh, A. K., Pareek, M., & Hanif, W. (2020). Is ethnicity linked to incidence or outcomes of covid-19?. British Medical Journal, 369.

Kwok, C. Y. T., Wong, M. S., Chan, K. L., Kwan, M. P., Nichol, J. E., Liu, C. H., ... & Kan, Z. (2021). Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19, a study in Hong Kong. Science of the Total Environment, 764, 144455.

Ladhani, S. N., Ireland, G., Baawuah, F., Beckmann, J., Okike, I. O., Ahmad, S., ... & Ramsay, M. E. (2021). Emergence of SARS-CoV-2 Alpha (B. 1.1. 7) variant, infection rates, antibody seroconversion and seroprevalence rates in secondary school students and staff: active prospective surveillance, December 2020 to March 2021, England. Journal of Infection, 83(5), 573-580.

Lai, P. C., Low, C. T., Tse,W. S.C., Tsui, C. K., Lee, H., Hui, P. K., 2013. Risk of tuberculosis in high-rise and high density dwellings: an exploratory spatial analysis. Environment Pollution. 183, 40–45.

Lai, K. Y., Webster, C., Kumari, S., & Sarkar, C. (2020). The nature of cities and the Covid-19 pandemic. Current Opinion in Environmental Sustainability, 46, 27-31.

Lei, H., Xu, X., Xiao, S., Wu, X., & Shu, Y. (2020). Household transmission of COVID-19-a systematic review and meta-analysis. Journal of Infection, 81(6), 979-997.

Li, Y., Qian, H., Hang, J., Chen, X., Cheng, P., Ling, H., ... & Kang, M. (2021). Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. Building and Environment, 196, 107788.

Lim, T., Cho, J., & Kim, B. S. (2010). The predictions of infection risk of indoor airborne transmission of diseases in high-rise hospitals: Tracer gas simulation. Energy and Buildings, 42(8), 1172-1181.

Lin, C., Lau, A. K., Fung, J. C., Guo, C., Chan, J. W., Yeung, D. W., ... & Lao, X. Q. (2020). A mechanism-based parameterisation scheme to investigate the association between transmission rate of COVID-19 and meteorological factors on plains in China. Science of The Total Environment, 737, 140348.

Ma, L., & Dill, J. (2015). Associations between the objective and perceived built environment and bicycling for transportation. Journal of Transport & Health, 2(2), 248-255.

Maas, J., Verheij, R. A., Groenewegen, P. P., De Vries, S., & Spreeuwenberg, P. (2006). Green space, urbanity, and health: how strong is the relation?. Journal of Epidemiology and Community Health, 60(7), 587-592.

Mao, J., & Gao, N. (2015). The airborne transmission of infection between flats in high-rise residential buildings: a review. Building and Environment, 94, 516-531.

Megahed, N. A., & Ghoneim, E. M. (2020). Antivirus-built environment: Lessons learned from Covid-19 pandemic. Sustainable Cities and Society, 61, 102350.

Ministry of Housing, Communities and Local Government. (2019). The English Indices of Deprivation. https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

Moghadam, S. T., Toniolo, J., Mutani, G., & Lombardi, P. (2018). A GIS-statistical approach for assessing built environment energy use at urban scale. Sustainable Cities and Society, 37, 70-84.

Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Science of the Total Environment, 728, 138884.

NHS Data Model and Dictionary. (2022). Middle Layer Super Output Area. https://www.datadictionary.nhs.uk/nhs_business_definitions/middle_layer_super_output_area

Pafka, E. (2020). As coronavirus forces us to keep our distance, city density matters less than internal density. Retrieved from The Conversation: https://the conversation.com/as-coronavirus-forces-us-to-keep-our-distance-city-density-matters-less-than-internal-density-137790.

Park, Neil (2021) Middle Super Output Area population estimates (supporting information). https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/middlesuperoutputareamidyearpopulationestimates

Paton, R.S., Overton, C.E., & Ward, T. (2022). The rapid replacement of the SARS-CoV-2 Delta variant by Omicron (B.1.1.529) in England. Science Translational Medicine, 14, eabo5395.

Public Health England. (2020). The Health Protections (Coronavirus, Restrictions) (England) Regulations 2020. https://www.legislation.gov.uk/uksi/2020/350/pdfs/uksi_20200350_en.pdf

Qu, G., Li, X., Hu, L., & Jiang, G. (2020). An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19). Environmental Science & Technology, 54 (7), 3730-3732

Raifman, M. A., & Raifman, J. R. (2020). Disparities in the population at risk of severe illness from COVID-19 by race/ethnicity and income. American Journal of Preventive Medicine, 59(1), 137-139.

Sallis, J. F., Adlakha, D., Oyeyemi, A., & Salvo, D. (2020). An international physical activity and public health research agenda to inform coronavirus disease-2019 policies and practices. Journal of Sport and Health Science, 9(4), 328.

Sharifi, A., & Khavarian-Garmsir, A. R. (2020). The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. Science of the Total Environment, 749, 142391.

Stier, A. J., Berman, M. G., & Bettencourt, L. (2020). COVID-19 attack rate increases with city size. arXiv preprint arXiv:2003.10376.

Tang, S., Xiao, Y., Yang, Y., Zhou, Y., Wu, J., & Ma, Z. (2010). Community-based measures for mitigating the 2009 H1N1 pandemic in China. PLoS one, 5(6), e10911.

Tong, H., & Kang, J. (2021). Characteristics of noise complaints and the associations with urban morphology: A comparison across densities. Environmental Research, 197, 111045.

Torjesen, I. (2021). Covid-19: Delta variant is now UK's most dominant strain and spreading through schools.

UK Health Security Agency. (2022). About the coronavirus (COVID-19) in the UK dashboard. https://coronavirus.data.gov.uk/about

UK Parliament. (2021). Coronavirus: A history of English lockdown laws. https://commonslibrary.parliament.uk/research-briefings/cbp-9068/

University College London. (2022). SPSS tutorial: Assumptions of Multiple Regression. https://www.open.ac.uk/socialsciences/spsstutorial/files/tutorials/assumptions.pdf

Wang, Y., Cooper, E., Tahmasebi, F., Taylor, J., Stamp, S., Symonds, P., ... & Mumovic, D. (2022). Improving indoor air quality and occupant health through smart control of windows and portable air purifiers in residential buildings. Building Services Engineering Research and Technology, 01436244221099482.

World Health Organization. (2022). WHO Coronavirus (COVID-19). Dashboard. https://covid19.who.int/

Xiao, H., Lin, X., Chowell, G., Huang, C., Gao, L., Chen, B., ... & Yang, H. (2014). Urban structure and the risk of influenza A (H1N1) outbreaks in municipal districts. Chinese Science Bulletin, 59(5), 554-562.

Xu, C., Wei, X., Liu, L., Su, L., Liu, W., Wang, Y., & Nielsen, P. V. (2020). Effects of personalized ventilation interventions on airborne infection risk and transmission between occupants. Building and Environment, 180, 107008.

Declaration of interests

☒ 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.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: