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.
Green spaces, especially nearby forest, may reduce the SARS-CoV-2 infection rate: A nationwide study in the United States

Bin Jiang\textsuperscript{a,b}, Yuwen Yang\textsuperscript{a,b}, Long Chen\textsuperscript{c}, Xueming Liu\textsuperscript{a,b}, Xueying Wu\textsuperscript{c}, Bin Chen\textsuperscript{d,e,f}, Chris Webster\textsuperscript{g}, William C. Sullivan\textsuperscript{h,i}, Linda Larsen\textsuperscript{j}, Jingjing Wang\textsuperscript{c}, Yi Lu\textsuperscript{k,*}

\textsuperscript{a} Urban Environments and Human Health Lab, HKUrbanLabs, Faculty of Architecture, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{b} Division of Landscape Architecture, Department of Architecture, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{c} Department of Architecture and Civil Engineering, College of Engineering, City University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{d} Future Urbanity & Sustainable Environment (FUSE) Lab, Division of Landscape Architecture, Department of Architecture, Faculty of Architecture, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{e} Urban Systems Institute, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{f} HKU Musketeers Foundation Institute of Data Science, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{g} HKUrbanLabs, Faculty of Architecture, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{h} HKU Musketeers Foundation Institute of Data Science, The University of Hong Kong, Hong Kong Special Administrative Region
\textsuperscript{i} Smart, Healthy Communities Initiative, University of Illinois at Urbana-Champaign, USA
\textsuperscript{j} Department of Architecture, University of Illinois at Urbana-Champaign, USA
\textsuperscript{k} Smart Energy Design Assistance Center, University of Illinois at Urbana-Champaign, USA

HIGHLIGHTS

- A greater dose of total green space tied to lower SARS-CoV-2 infection rate.
- A greater dose of forest was associated with a lower infection rate.
- Forest outside park yielded a greater beneficial effect than forest inside park.
- Open space variables yielded mixed associations with the infection rate.
- Optimal buffer distances of forest inside and outside park are found.

ABSTRACT

The coronavirus pandemic is an ongoing global crisis that has profoundly harmed public health. Although studies found exposure to green spaces can provide multiple health benefits, the relationship between exposure to green spaces and the SARS-CoV-2 infection rate is unclear. This is a critical knowledge gap for research and practice. In this study, we examined the relationship between total green space, seven types of green space, and a year of SARS-CoV-2 infection data across 3,108 counties in the contiguous United States, after controlling for spatial autocorrelation and multiple types of covariates. First, we examined the association between total green space and SARS-CoV-2 infection rate. Next, we examined the association between different types of green space and SARS-CoV-2 infection rate. Then, we examined forest–infection rate association across five time periods and five urbanicity levels. Lastly, we examined the association between infection rate and population-weighted exposure to forest at varying buffer distances (100 m to 4 km). We found that total green space was negative associated with the SARS-CoV-2 infection rate. Furthermore, two forest variables (forest outside park and forest inside park) had the strongest negative association with the infection rate, while open space variables had mixed associations with the infection rate. Forest outside park was more effective than forest inside park. The optimal buffer distances associated with lowest infection rate are within 1,200 m for forest outside park and within 600 m for forest inside park.
1. Introduction

The SARS-CoV-2 pandemic has profoundly affected people’s health and well-being globally (World Health Organization, 2021). Studies have explored associations between SARS-CoV-2 infection rate and numerous social, economic, and medical factors (Badr et al., 2020; Carteni et al., 2020; Mena et al., 2021; Muller et al., 2021). Fewer studies have explored the relationship between the built environment and infection rate. We know that architectural and landscape elements, especially green spaces, profoundly impact people’s physical and mental health and well-being. And yet, we have largely overlooked the relationship between green spaces and SARS-CoV-2 infection rate (Frumkin, 2021).

Several recent studies on green space and SARS-CoV-2 infection rate in the U.S. suggest that green spaces may provide some protection against infection (Klompmaker et al., 2021; Lu et al., 2021). However, these preliminary studies were conducted before the much larger second and third waves of the pandemic hit the U.S. Furthermore, these studies used simple measures of greenness, such as the normalized difference vegetation index (NDVI) and leaf area index (LAI). Though effective at capturing the amount of greenness in an area, these measures fail to consider factors that are likely to impact the relationship between SARS-CoV-2 infection and green settings. Previous studies suggest factors such as green space type (Apinan et al., 2016; Allard-Poesi et al., 2022), population density (Stier et al., 2020), distance to green settings (Ekkel & de Vries, 2017; Kim & Miller, 2019) exert significant impacts on health outcomes. Among all factors, park has been frequently reported as an environmental factor that has a significant relationship with the SARS-CoV-2 infection rate (Johnson et al., 2020; Ma et al., 2022).

The objective of this study is to examine the links between various types of green spaces and SARS-CoV-2 infection rate for an entire year after controlling for all covariates. Without understanding these relationships, we may lose the opportunity to build supportive environments that increase our ability to resist infection, especially in our most vulnerable communities.

1.1. Accumulating evidence: The health benefits of exposure to green spaces

There is overwhelming evidence documenting the health benefits of green space exposure at national (Lu, Chen, et al., 2021; Nowak et al., 2014), city (Donovan et al., 2011), and community levels (Chang et al., 2021; Kuo & Sullivan, 2001). Contact with green spaces is associated with improved mental (Jiang et al., 2014; Jiang et al., 2018; He et al., 2022) and physical health outcomes (Lu, 2018; Mitchell & Popham, 2008), and the effects are complex and interdependent (Jiang et al., 2015; Sullivan & Bartlett, 2005). For example, green spaces significantly reduce mental stress and fatigue (Jiang et al., 2020; Jiang et al., 2021), which positively influence immune function and promote physical health (Kuo, 2015; Parsons et al., 1998). Many studies have found that green open spaces are negatively associated with chronic health outcomes because they facilitate physical activity and social interactions, reduce air pollution, and enhance immune function (Jiang et al., 2014; Kuo, 2015).

However, we still know little about whether green spaces impact SARS-CoV-2 infection rate. Recent studies have generated mixed results. Several studies identified negative associations between greenness and SARS-CoV-2 infection rate (Klompmaker et al., 2021; Spotswood et al., 2021). However, one study found that highly connected green spaces were associated with higher risk of SARS-CoV-2 transmission (Pan et al., 2021). Another study pointed out that although outdoor transmission of SARS-CoV-2 is less common, the risk does exist (Bulfone et al., 2021).

1.2. Necessity to compare health benefits of green spaces inside and outside park

Park is a mixture of many types of green spaces, such as open lawn, forest, and shrubs. As a critical type of public space for recreativist and social activities, parks may have greater impacts on public health and well-being than private green spaces (Ventor et al., 2020, 2021). And yet, studies exploring the association between parks and SARS-CoV-2 infection rate reveal mixed results. One study found that park use had no impact on infection rate (Kartal et al., 2021), while another study found that the availability of parks was associated with lower risk of SARS-CoV-2 transmission (Wang et al., 2021). Yet another study found park use decreased pre-peak SARS-CoV-2 infection rate (Johnson et al., 2020). Hitherto, we are not clear whether the relationships between different types of green spaces and SARS-CoV-2 infection rate are the same; or whether the effect of green spaces inside park is significantly different from green space outside park.

1.3. A critical knowledge gap: The relationship between green spaces and SARS-CoV-2 infection rate

While it is widely recognized that green spaces have a significant positive effect on human health, we know much less about how exposure to green spaces impacts infectious diseases such as SARS-CoV-2. One study examined the 135 most urbanized counties in the contiguous United States and found that a higher ratio of green space was significantly associated with lower racial disparity in SARS-CoV-2 infection rate at the county level (Lu et al., 2021). However, this study considered a relatively small number of highly urbanized counties (135) and used infection data from January through June 2020, a relatively short time period given the length of the pandemic.

Another study examined the association between county-level NDVI and SARS-CoV-2 infection and death rate for 2297 counties in the United States (Klompmaker et al., 2021). The study reported greenness was negatively associated with county-level SARS-CoV-2 infection rate. However, this study also used data from a relatively short time period (March to June 2020) and did not identify the types of green spaces that were associated with lower infection rate. The study also failed to account for important covariates, including transportation infrastructure and services (Carteni et al., 2020; Tiarachi & Cats, 2020), political and administration factors (Clinton et al., 2021), human mobility (Muller et al., 2021), commuting mode (Figueroa et al., 2021), and employment status (Mena et al., 2021).

There is an urgent need for a more comprehensive assessment of the relationship between green spaces and SARS-CoV-2 infection rate. Understanding this relationship will enable planners and designers to develop appropriate environmental interventions that reduce the risk of infection for current and future airborne infectious diseases.

1.4. Research questions

In this study, we asked four layers of questions: 1) What is the association between total green space and SARS-CoV-2 infection rate? 2) What are the associations between various types of green spaces and SARS-CoV-2 infection rate and what are key green spaces that have the greatest impacts on the associations? 3) What are the relationships between key green spaces and SARS-CoV-2 infection rate across various levels of urbanicity and over distinct time periods of SARS-CoV-2 pandemic? 4) What are the optimal buffer distances of key green
spaces exposure associated with reduced levels of SARS-CoV-2 infection rate?

2. Material and methods

2.1. Study design

We investigated the association between total green space, various types of green space, and SARS-CoV-2 infection rate in the contiguous United States from January 22 to December 31, 2020. We also examined associations between forest and SARS-CoV-2 infection rate in counties across five levels of urbanicity and five different time periods. To identify the optimal buffer distance, we examined the relationship between population-weighted exposure to forest within various buffer distances (100 m to 4 km) from human population distribution and the SARS-CoV-2 infection rate. We used the county as the basic unit of analysis and included a total of 3,108 counties in the contiguous United States.

2.1.1. SARS-CoV-2 infection

We obtained the number of positive cases of SARS-CoV-2 infection rate from January 22, 2020 to December 31, 2020 from the Centers for Disease Control and Prevention (CDC) and state- and local-level public health agencies (USAFACTS, 2021). We chose December 31 as the endpoint as the SARS-CoV-2 pandemic neared its peak at approximately this week, despite some fluctuations (Fig. 1). Moreover, this endpoint fell prior to the rollout of large-scale vaccination programs and the change in the U.S. presidency, allowing us to avoid the confounding effects of political and programmatic factors. The research period was divided into five periods to investigate the temporal association between SARS-CoV-2 infection rate and green spaces. We select the break point of the five time periods based on the development stage of the SARS-CoV-2 infection in 2020, as the severity of SARS-CoV-2 infection rate may affect social distancing policies and people’s mobility patterns including green space usage during the pandemic (Heo et al., 2020; Tokey, 2021). As shown in Fig. 1, period 1 included the onset and early outbreak of SARS-CoV-2 (January 22 to March 30); period 2 included the first wave of SARS-CoV-2 infection in 2020 (March 31 to June 7), period 3 included the second wave in 2020 (June 8 to August 15), period 4 included a stagnation period in 2020 (August 16 to October 23), and period 5 included the peak in 2020 (October 24 to December 31, 2020).

2.1.2. Green spaces

We quantified the total and seven types of green space with predominant natural elements assessed at a 30-m resolution using National Land Cover Datasets (NLCD, 2016): open space inside park, open space outside park, forest inside park, forest outside park, shrub and scrub, herbaceous, hay and pasture (Table S1). County-level open space and forest were divided into factors within and outside parks using the USA Parks dataset from Esri (Esri, 2021). The ratio of total and seven types of green space were measured as the area of total and each type of green space within a county divided by the total county area (Fig. 2). We also calculated population-weighted exposure to forest within varying buffer sizes (100 - 4 km) in each county in Google Earth Engine (GEE). We extracted the forest inside park and forest outside park using the NLCD 2016 and USA Parks boundary (Esri, 2021). Then, we located the population distribution of the residents in the contiguous United States from the WorldPop Global Project Population Data 2020 (Sorichetta et al., 2015). The WorldPop Global Project Population Data 2020 estimated the number of people residing in each 100x100m grid cell matched to their associated administrative units. The original 30 m NLCD 2016 land cover map was reprojected to a 100 m resolution to match the population data. The population-weighted exposure to forest within various buffer distances in each county was calculated using the following Eq. (1) (Chen et al., 2018),

\[
FE = \frac{\sum_{i=1}^{N} P_i \times F_i}{\sum_{i=1}^{N} P_i}
\]

where \(P_i\) represents the population of the \(i\)th grid, \(F_i\) represents the forest cover of the \(i\)th grid with a buffer size of \(b\) meters, \(N\) denotes the total number of grids for a given county, and \(FE\) is the estimated forest exposure area per person for the given county.

The population-weighted exposure considers population spatial distribution in forest exposure estimates by giving proportionally greater weight to forest near densely populated areas. The forest exposure received by the population living in a grid is not only the forest cover within this grid, but also includes the forest around the living grid in a certain spatial range (e.g., 200 m, 400 m) (Fig. 3). In this calculation, the buffer distance is calculated from the center of the grid, a grid is included in the buffer zone if the center point of the grid falls inside the dashed-line circle. We select the 4 km threshold because most walking trips are within 4 km (Yang & Diez-Roux, 2012). We set the buffer intervals of 200 m for buffer distances less than or equal to 2 km and 500 m for buffer distances 2 to 4 km.

2.1.3. Levels of urbanicity

We categorized all counties into five levels of urbanicity based on the 2013 Urban-Rural Classification Scheme from the National Center for Health Statistics, which is well-suited for health analyses (NCHS, 2017). Urbanicity level 1 is the most urbanized and level 5 is the most rural (Fig. 4 & Table S2). Previous studies have identified an urban–rural disparity in the prevalence of SARS-CoV-2 infection in the United States (Huang et al., 2021; Pro et al., 2020). A recent review study also found exposure to green space has a heterogeneous effect on health outcomes across different levels of urbanicity (Browning et al., 2022).

2.1.4. Covariates

Additionally, we adjusted for potential covariates which significantly impact SARS-CoV-2 infection rates, including healthcare and testing rate (Wu et al., 2020), pre-existing chronic diseases (i.e., hypertension and diabetes) (Clerkin et al., 2020; Fang et al., 2020; Sattar et al., 2020), socioeconomic and demographic factors (i.e., racial minority and elderly) (Abedi et al., 2020; Clouston et al., 2021; Karaye & Horney, 2020), policies and policy factors (i.e., political affiliation and stay-at-home orders) (Neelon et al., 2021; Fowler et al., 2021), behavioral factors (i.e., mobility pattern and social distancing) (Badr et al., 2020; McGrail et al., 2020), and environmental factors (i.e., air pollution, crowded housing, and airport density) (Chakrabarty et al., 2021; Gaskin et al., 2021; McLaughlin et al., 2020). The definition of all covariates is presented in Table S2.

2.1.5. Descriptive statistics of variables

The descriptive statistics for each variable entered in the models are presented in Table 1 (See Table S3 & S4 for the descriptive information for all variables). The correlation matrix of all variables in the final model is given in Supplementary Fig. 1.

2.2. Statistical analysis

This study evaluates the association between green space and SARS-CoV-2 infection rate, with negative binomial mixed-effect models. The negative binomial model fits our overdispersion of the dependent variable, i.e., infection rate. We confirmed the spatial autocorrelation presence using Moran’s I test. The mixed-effect model can account for the clustering in the data of counties from the same state. We included a random effect of state to account for the non-independence of county-level data. The intra-class correlation coefficient (ICC) was 0.67, indicating that a 67% variation in county-level infection rate was attributed to the clustering structure of our data. It also supports the mixed-effect model is necessary. In all models, all explanatory variables were
Fig. 1. County-level Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infection (cases per 100,000 population) in total and across the five subperiods. (A) The overall research period, from January 22 to December 31, 2020. (B) Period 1: January 22 to March 30, 2020. (C) Period 2: March 31 to June 7, 2020. (D) Period 3: June 8 to August 15, 2020. Period 4: August 16 to October 23, 2020. (F) Period 5: October 24 to December 31, 2020. (G) Diagram of cases of SARS-CoV-2 from February 2020 to July 2021 in the United States. The SARS-CoV-2 pandemic emerged out the U.S. in 2020 and comprised several different periods, with low, moderate, and high infection rates. The infection rate in 2021 was significantly attenuated due to wide-scale vaccination, so this data was not included.
Fig. 2. Ratio of seven types of green spaces at a county level. Values represent the (A) forest inside park, (B) forest outside park, (C) shrub/scrub, (D) grassland/herbaceous, (E) open space inside park, (F) open space outside park, (G) hay/pasture at the county level, calculated as the total area of each green space divided by the county area. Data were extracted from the NLCD landcover dataset.
centered and scaled; the covariates were adjusted including healthcare and testing factors, pre-existing chronic disease factors, socioeconomic and demographic factors, politics and policy factors, behavioral factors, and environmental factors.

We conduct two sets of statistical analyses. In the first set, Model 1 estimates the effect of total green space on the infection rate. Model 2 estimates the effects of seven types of green space. In the second set, Model 3 estimates the effects of forest in five time periods and five urbanicity levels using a series of separate models. The testing and regulation and policy variables (i.e., public mask mandate and stay-at-home order) were calculated in a time-sensitive way to reflect the variation in each period of time (see calculation method in Supplementary Table S2). Model 4 examines the dose–response effect of population-weighted exposure to forest at various buffer distances. This allows us to identify the optimal exposure distances from forest. In all models, we reported the effect size (β) and/or the incidence rate ratio (IRR), which compares the effect of independent variables on the SARS-CoV-2 infection rate. The IRR estimates the estimated rate ratio of SARS-CoV-2 infection rate for a one-unit increase in change of a variable, given the other variables are held constant in the model.
Table 1
Descriptive data for SARS-CoV-2 infection rate, socioeconomic and demographic, healthcare and testing, pre-existing chronic disease, policy and regulation, behavioral, environmental, and green space factors.

| Variable Categories                          | Variables                                      | Min     | Max     | Mean    | SD      | Unit or Formula |
|----------------------------------------------|------------------------------------------------|---------|---------|---------|---------|-----------------|
| SARS-CoV-2 Infection rate                    | Infection rate                                 | 420.885 | 2717.734| 6570.950| 2762.664| Cases per 100 k |
| Socio-economic and Demographic factors       | Population density                             | 0.094   | 19625.842| 98.221  | 559.972 | Persons per km² |
|                                              | Black non-Hispanic                             | 0.000   | 87.400   | 8.957   | 14.520  | Ratio           |
|                                              | Population aged 65 above                       | 3.800   | 55.600   | 18.428  | 4.542   | Ratio           |
|                                              | Gini Index                                     | 0.257   | 0.665    | 0.446   | 0.036   | Range 0–1       |
|                                              | Median home value                              | 21000.00| 1057000.00| 146157.40| 89047.61| USD²           |
|                                              | Unemployment rate                              | 1.300   | 18.100   | 4.093   | 1.397   | Ratio           |
|                                              | Population without high school diploma         | 1.200   | 66.300   | 13.448  | 6.342   | Ratio           |
| Healthcare and testing factors               | Population without insurance                   | 2.300   | 33.700   | 11.418  | 5.106   | Ratio           |
|                                              | COVID-19 testing rate                          | 0.000   | 1711.099 | 170.197 | 107.693 | Per 100 k       |
| Pre-existing chronic disease factors         | Diabetes rate                                  | 2.200   | 28.700   | 10.506  | 3.526   | Ratio           |
|                                              | Obesity rate                                   | 12.30   | 57.90    | 32.76   | 5.68    | Ratio           |
|                                              | Stroke mortality                               | 14.000  | 92.500   | 39.776  | 8.153   | Per 100 k       |
|                                              | Hypertension mortality                         | 19.300  | 587.300  | 131.825 | 55.366  | Per 100 k       |
|                                              | Heart failure mortality                        | 19.400  | 304.300  | 108.514 | 25.488  | Per 100 k       |
| Politics and policy factors                  | State Governor Party                           | 0.000   | 1.000    | 0.435   | 0.496   | Democratic/Republican |
|                                              | Stay-at-home orders                            | 3.856   | 6.897    | 4.441   | 0.447   | Range 1–7       |
|                                              | Public mask mandates                           | 0.000   | 1.000    | 0.497   | 0.339   | Yes/No          |
|                                              | Business closing and reopening                 | 4.781   | 8.233    | 6.331   | 0.683   | Range 1–5       |
| Behavioral factors                           | Smoker                                         | 5.909   | 41.491   | 17.446  | 3.554   | Ratio           |
|                                              | Essential worker                               | 0.178   | 0.791    | 0.526   | 0.073   | Ratio           |
|                                              | Foot traffic to all points of interest (POI)   | 0.000   | 10.507   | 1.505   | 0.904   | Per person      |
|                                              | Commute to work by walking or bicycle          | 0.000   | 42.410   | 3.427   | 3.146   | Ratio           |
|                                              | Leisure time physical inactivity               | 9.400   | 49.800   | 26.238  | 5.498   | Ratio           |
|                                              | Mobility                                       | 0.409   | 553.723  | 8.118   | 14.190  | Km              |
|                                              | Normalized mobility index                      | 17.156  | 947.757  | 79.520  | 40.596  | Percentage      |
| Environmental problems                       | Severe housing problems                       | 2.700   | 39.100   | 14.308  | 4.338   | Ratio           |
|                                              | Overcrowded housing                            | 0.000   | 16.900   | 2.314   | 1.816   | Ratio           |
|                                              | Proximity to highway                           | 0.000   | 16.400   | 1.943   | 1.871   | Ratio           |
|                                              | Airport density                                | 0.000   | 0.059    | 0.002   | 0.002   | Number per km²  |
|                                              | Railway density                                | 0.000   | 2.625    | 0.062   | 0.111   | Length of km per km² |
|                                              | Highway and secondary road density             | 0.000   | 1.577    | 0.105   | 0.132   | Length of km per km² |
|                                              | PM₂.⁵                                         | 1.500   | 16.000   | 7.640   | 1.674   | µg/m³           |
|                                              | PM₁₀                                          | 7.476   | 57.922   | 17.417  | 4.755   | µg/m³           |
|                                              | NO₂                                           | 2.896   | 27.402   | 13.321  | 3.337   | Ppb             |
|                                              | Average temperature                            | 36.899  | 79.218   | 57.687  | 7.941   | Degrees Fahrenheit |
|                                              | Wind speed                                     | 3.969   | 9.923    | 7.026   | 0.800   | m/s             |
| Green space factors                          | Total green space                              | 0.008   | 0.998    | 0.622   | 0.291   | ratio           |
|                                              | Shrub/scrub                                    | 0.000   | 0.976    | 0.085   | 0.181   | ratio           |
|                                              | Grassland/herbaceous                           | 0.000   | 0.977    | 0.094   | 0.169   | ratio           |
|                                              | Hay/pasture                                    | 0.000   | 0.799    | 0.101   | 0.123   | ratio           |
|                                              | Open space inside park                         | 0.000   | 0.116    | 0.002   | 0.007   | ratio           |
|                                              | Open space outside park                       | 0.000   | 0.279    | 0.040   | 0.034   | ratio           |
|                                              | Forest inside park                             | 0.000   | 0.899    | 0.056   | 0.129   | ratio           |
|                                              | Forest outside park                            | 0.000   | 0.905    | 0.244   | 0.218   | ratio           |

The variance inflation factor (VIF) test was used to identify multi-collinearity between the independent variables, and variables with a VIF ≥ 4 were excluded from our models (O’Brien, 2007). All analyses were performed in R v.4.1.2 (R Core Team, 2020). Moran’s I test was performed using the package ‘spdep’ (Bivand & Wong, 2018) and negative binomial mixed effect models were performed using the package “lme4” (Bates et al., 2014).

3. Results

3.1. Associations of total green space and various types of green space with SARS-CoV-2 infection rate

The results of the negative binomial mixed effect model for total green space, seven types of green space and SARS-CoV-2 infection rate are shown in Tables 2 & 3, respectively. After controlling for covariates, the total green space has a significant and negative association with the SARS-CoV-2 infection rate ($\beta = -0.059, p < 0.0001$) (Fig. 5). For the seven types of green spaces, forest inside park ($p < 0.0001$), forest outside park ($p < 0.0001$), and hay/pasture ($p < 0.001$) are significantly negatively associated with SARS-CoV-2 infection rate. Open space inside park is positively associated with SARS-CoV-2 infection rate ($p < 0.001$) (Fig. 6).

Among all green spaces, forest outside park and forest inside park have the greatest effect size on SARS-CoV-2 infection rate ($\beta = -0.087$ and $\beta = -0.058$, respectively). We found that a one-unit increase in forest inside park is associated with a 5.6% decrease in SARS-CoV-2 infection rate (IRR 95% CI: 0.95%–7.5%), and a one-unit increase in forest outside park is associated with an 8.3% decrease in SARS-CoV-2 infection rate (IRR 95% CI: 7.0%–9.7%) (Table 4).

3.2. Forest-SARS-CoV-2 infection rate associations across five levels of urbanicity

After identifying forest outside park and forest inside park as key green spaces, we further ran separate negative binomial mixed effect models to measure associations between the forest variables and SARS-CoV-2 infection rate across five levels of urbanicity (Fig. 7). The descriptive data of all variables across five urbanicity levels are presented in Table S4 and results are presented in Table S5. Forest inside park is significantly negatively associated with infection rate at urbanicity levels 3 and 5; the association is strongest at level 3 ($\beta = -0.097, p$
Forest outside park is significantly negatively associated with infection rate across urbanicity levels 2 to 5; the association is strongest at level 5 ($\beta = -0.095$, $p < 0.0001$).

### 3.3. Forest-SARS-CoV-2 infection rate associations across five time periods

Next, we examined the associations between the two forest variables and SARS-CoV-2 infection rate across five time periods. Fig. 7 and Table S6 show the effect sizes of two forest variables on SARS-CoV-2 infection rate during different time periods. Forest outside park has a significant and negative association with the infection rate from periods 2 to 5, and the association is strongest in period 2 ($\beta = -0.174$, $p = 0.0001$). Forest inside park has a significant and negative association in time periods 2, 4 and 5, with the largest effect size in period 2 ($\beta = -0.149$, $p < 0.0001$).

### 3.4. Associations of population-weighted exposure to forest at varying buffer distances with SARS-CoV-2 infection rate

Next, we examined the association of population-weighted exposure to forest at various buffer sizes within walking distance (100 m to 4 km) with the SARS-CoV-2 infection rate. We report the standardized coefficient value representing effect sizes and the 95 % CI in Table 5 and Fig. 8.

Forest inside park is significantly negatively associated with SARS-CoV-2 infection rate from 200 m to 4 km and reaches an optimal effect at 4,000 m. The exposure buffer-response curve shows the effect size of forest inside park increases as buffer size increases to 600 m, then decreases beyond 600 m, and increase again at 2,500 m, though the increase in effect size remains limited (600 m: $\beta = -0.023$ vs 4,000 m: $\beta = -0.026$). As shown in Table S13, per one unit increase in exposure to forest in park is linked to 2.3 % decrease in infection rate within 600 m buffers (IRR 95 % CI: 0.6 %–4%) and a 2.6 % decrease at 4,000 m buffer (IRR 95 % CI: 0.9 %–4.2 %) (Table 6).

Exposure to forest outside park is significantly negatively associated with SARS-CoV-2 infection rate from 100 m to 4 km, reaching the largest effect size around 1,200 m. The exposure buffer-response curve suggests the effect size of forest outside park increases between 100 m and 1,200 m, then decreases beyond 1,200 m, and increase again at 2,500 m, though the increase in effect size remains limited (600 m: $\beta = -0.023$ vs 4,000 m: $\beta = -0.026$). As shown in Table S13, per one unit increase in exposure to forest outside park is linked to 3.2 % decrease in infection rate within 600 m buffers (IRR 95 % CI: 0.6 %–4%) and a 2.6 % decrease at 4,000 m buffer (IRR 95 % CI: 0.9 %–4.2 %) (Table 6).

### 4. Discussion

This study examines the relationship between green spaces and SARS-CoV-2 infection rate across all 3,108 counties during 2020 in the contiguous United States after controlling for multiple covariates. In the following sections, we first provide interpretations of key findings. Then, we discuss potential impacts and implications of this study. Lastly, we discuss the limitations of this study and opportunities for future research.

#### 4.1. Interpretation of key findings

##### 4.1.1. Why total green space has a significant and negative association with the SARS-CoV-2 infection rate?

We found that the total green space has a significant negative
association with the infection rate. Although many studies have summarized that green spaces can make positive impacts on mental and physical health (Jiang et al., 2015; Labib et al., 2020; Markveych et al., 2017; Zhang, Yu, Zhao, Sun, & Vejre, 2020), a comprehensive interpretation on why green spaces have positive impacts on SARS-CoV-2 infection rate is rarely presented. We argue the association can be interpreted by four causal mechanisms.

Being in green spaces allows people to have normal recreational and social life while maintaining a safer social distance than being in indoor spaces. The primary pathway of SARS-CoV-2 transmission is via aerosol particles and droplets that are exhaled by human hosts (Bourouiba, 2020; Klompas et al., 2020; Zhang, Li, et al., 2020). Hence, virus transmission is less likely to occur outdoors than indoors (Leclerc et al., 2020). Comparing to indoor spaces, green spaces can also enable people to maintain adequate physical distancing (Leclerc et al., 2020). In other words, green spaces are a relatively safer social environment than indoor spaces during the SARS-CoV-2 pandemic: they encourage people to leave indoor environments and participate in outdoor activities with other people while maintaining safe social distances (Lu et al., 2021; Schipperijn et al., 2013; Sullivan et al., 2004). Green spaces not only invite people to go outdoors more often (Coley et al., 1997), but also encourage them to stay outdoors longer, reducing their time spent indoors (Braubach et al., 2017; Coley et al., 1997; Grahm & Stigsdotter, 2003). Nevertheless, it is important to point out the “safer social distance” is not always “safe enough”. The risk of infection might be high if large gathering occur in green spaces or they have close contact during social activities, such as conversation, dance party, food-sharing party (Domenech-Montoliu et al., 2021; Peng et al., 2022).

Green spaces can promote physical activity thus enhance immune functioning. Previous studies have reported that green spaces promote physical activity (Cohen et al., 2007; Lu et al., 2017; Lu et al., 2021). Exercising while viewing green landscapes can produce synergic health benefits. Several studies have suggested that green spaces could reduce the risk of obesity by promoting physical activity, making people less vulnerable to SARS-CoV-2 infection (Jia et al., 2021; Jordan & Adab, 2020; Sattar et al., 2020). Physical activities in green spaces can enhance immune functioning (Li et al., 2010; Amatriain-Fernández et al., 2020), thus strengthening resistance to SARS-CoV-2 infection.

Green spaces can reduce mental stress thus enhance immune functioning. Adverse mental state can make people more vulnerable to SARS-CoV-2 virus (Qin et al., 2020; Wang et al., 2021; Yang et al., 2020). Extensive evidence suggests viewing or being in a forest can reduce mental stress (Hunter et al., 2017; Jiang et al., 2016; Ulrich et al., 1991). Elevated stress levels were found to weaken immune functioning (Dhabhar, 2014; Marketon & Glaser, 2008). Exposure to forest enhance immune functioning by increasing the numbers of Natural Killer (NK) cells, lymphocytes, and enhances human NK activity, which would also

| Variable Categories          | Variables        | Coefficient | SE    | Z Value | p Value |
|------------------------------|------------------|-------------|-------|---------|---------|
| Socioeconomic and demographic factors | Population density | 0.002 | 0.007 | 0.242 | 0.809 |
|                              | Black non-Hispanic | -0.019 | 0.009 | -2.196 | 0.028 |
|                              | Population aged 65 above | -0.069 | 0.007 | -10.430 | <0.0001*** |
|                              | Gini Index | 0.027 | 0.006 | 4.375 | <0.0001*** |
|                              | Median home value | -0.033 | 0.010 | -3.332 | 0.001** |
|                              | Unemployment rate | -0.016 | 0.007 | -2.209 | 0.027 |
|                              | Population without high school diploma | 0.078 | 0.010 | 8.085 | <0.0001*** |
| Healthcare and testing factors | Population without insurance | -0.012 | 0.013 | -0.868 | 0.385 |
|                              | COVID-19 testing rate | 0.131 | 0.007 | 19.912 | <0.0001*** |
| Pre-existing chronic disease factors | Diabetes rate | 0.004 | 0.006 | 0.697 | 0.486 |
|                              | Obesity rate | 0.000 | 0.007 | -0.052 | 0.959 |
|                              | Stroke mortality | -0.004 | 0.007 | -0.570 | 0.569 |
|                              | Hypertension mortality | 0.005 | 0.007 | 0.753 | 0.451 |
|                              | Heart disease mortality | 0.011 | 0.007 | 1.706 | 0.088 |
| Behavioral factors           | Smoker | -0.013 | 0.011 | -1.220 | 0.223 |
|                              | Essential worker | 0.040 | 0.008 | 4.977 | <0.0001*** |
|                              | Foot traffic to all POI | 0.032 | 0.006 | 5.650 | <0.0001*** |
|                              | Commute to work by walking or bicycle | -0.029 | 0.006 | -4.689 | <0.0001*** |
|                              | Leisure time physical inactivity | 0.017 | 0.007 | 2.319 | 0.020 |
|                              | Mortality | -0.003 | 0.005 | -0.597 | 0.551 |
|                              | Normalized mobility index | 0.005 | 0.005 | 0.980 | 0.327 |
| Politics and policy factors  | State governor party | -0.039 | 0.063 | -0.614 | 0.539 |
|                              | Stay-at-home orders | 0.012 | 0.010 | 1.170 | 0.242 |
|                              | Public mask mandates | -0.105 | 0.058 | -1.815 | 0.070 |
|                              | Business closing and reopening | 0.000 | 0.006 | 0.024 | 0.983 |
| Environmental factors        | Severe housing problem | -0.024 | 0.008 | -2.969 | 0.003* |
|                              | Overcrowded housing | 0.033 | 0.008 | 4.379 | <0.0001*** |
|                              | Proximity to highway | 0.016 | 0.006 | 2.864 | 0.004** |
|                              | Airport density | -0.031 | 0.005 | -5.654 | <0.0001*** |
|                              | Railway density | -0.013 | 0.006 | -2.079 | 0.038 |
|                              | Highway and secondary road density | -0.009 | 0.009 | -0.968 | 0.333 |
|                              | PM 2.5 | 0.043 | 0.009 | 4.909 | <0.0001*** |
|                              | PM 10 | 0.023 | 0.011 | 2.032 | 0.042 |
|                              | NO2 | 0.016 | 0.009 | 1.876 | 0.061 |
|                              | Average temperature | -0.109 | 0.015 | -7.105 | <0.0001*** |
|                              | Wind speed | 0.006 | 0.007 | 0.881 | 0.378 |
| Green space factors          | Shrub/Scrub | 0.019 | 0.009 | 2.262 | 0.024 |
|                              | Grassland/ Herbaceous | 0.017 | 0.007 | 2.262 | 0.024 |
|                              | Hay/Pasture | -0.022 | 0.006 | -3.511 | 0.000** |
|                              | Open space inside park | 0.021 | 0.008 | 2.724 | 0.006* |
|                              | Open space outside park | 0.016 | 0.008 | 2.040 | 0.041 |
|                              | Forest inside park | -0.058 | 0.009 | -6.726 | <0.0001*** |
|                              | Forest outside park | -0.087 | 0.008 | -11.466 | <0.0001*** |

Note: * indicates p < 0.01; ** indicates p < 0.001; *** indicates p < 0.0001.
strengthen resistance to SARS-CoV-2 infection (Li, 2010; Li et al., 2007). Green spaces can remove ambient pollutants to reduce transmission of virus. Green spaces in urban and rural areas can improve air quality by removing particulates and absorbing aerosols through leaf stomata (Janhäll, 2015; Kumar et al., 2019; Nowak et al., 2006, Nowak et al., 2013, Nowak et al., 2014, Nowak et al., 2018). Because SARS-CoV-2 is transmitted through particles and aerosols, these green spaces may reduce SARS-CoV-2 infection risk (Bourouiba, 2020). Those who have daily exposure to green spaces are less vulnerable to SARS-CoV-2 infection (Fattorini & Regoli, 2020; Paital & Agrawal, 2020; Zhu et al., 2020).

4.1.2. Why do forests have a stronger negative association with infection rate than other green spaces?

In this study, “forest” was defined as “an area dominated by trees, generally greater than 5 m tall, and accounting for greater than 20 % of total vegetation cover” (NLCD, 2016). Forests included large natural forests in rural areas and moderate or smaller patches of trees in urban, suburban, and rural areas.

We found forests have a stronger tie with lower SARS-CoV-2 infection rate than other types of green space. We propose three reasons for that difference. First, forests are more likely to entice people outdoors than other types of green spaces. During the pandemic, more people chose to visit natural forest parks than urban parks for recreational and exercise (Lu et al., 2021). Forests create a more comfortable microclimate than other green spaces without large tree canopy (Li et al., 2019; Ziter et al., 2019). Forested areas, which in our study include lawn or grassland partially covered by tree canopy, provide a more comfortable environment for outdoor activities than a lawn or park without shade.

Second, forests are more effective to boost immune function than other green spaces because forests are more mentally restorative than other green spaces (Li, 2010; Li et al., 2007; Lyu et al., 2019). Forests have a stronger presence of “positive distractions” than other green spaces (Beil & Hanes, 2013; Jiang, Wang, et al., 2019; Ulrich, 1997): Forests have a complex three-dimensional profile with trees of different sizes, type and age, and a diversity of vegetation, animals, and insects (Beil & Hanes, 2013; Brockerhoff et al., 2017; Trochta et al., 2017; Wood et al., 2018). Those positive distractions allow people to have a greater...
level of stress reduction, thus enhance their immune functioning to resist the risk of infection (Morey et al., 2015).

Lastly, forests are also more effective at reducing air pollutants than other green spaces, which makes forests more effective to reduce risk of SARS-CoV-2 virus transmission. Researchers found that forests capture particulate pollutants more efficiently than grassland and shrubs, as forests often have more complex foliage, a larger number of trees, more diverse tree heights and canopy sizes, and more diverse tree species (Beckett et al., 2000). A nationwide study in the contiguous United States found that trees and forests removed approximately 17.4 million tonnes of air pollution in 2010, which was estimated to lead to 850 fewer deaths and 670,000 fewer incidents of acute respiratory symptoms (Nowak et al., 2006).

Fig. 6. Relationships between different types of green spaces, covariates, and SARS-CoV-2 infection (Model 2). Coefficient values represent effect sizes from the negative binomial mixed effects model for the relationship between infection rate of SARS-CoV-2 (cases per 100,000 people) and all variables. Bars represent 95% CIs and significant variables are shown out red, orange, and yellow. Note: * p < 0.01; ** p < 0.001; *** p < 0.0001.

Table 4
Incident rate ratio of seven types of green spaces with SARS-CoV-2 infection rate in overall model after adjusting for covariates (Model 2).

| Green spaces                      | SARS-CoV-2 IRR (95% CI)  |
|----------------------------------|-------------------------|
| Shrub/Scrub                      | 1.0195 (1.0026, 1.0367) |
| Grassland/Herbaceous             | 1.0167 (1.0022, 1.0314) |
| Hay/Pasture                      | 0.9785 (0.9668, 0.9905) |
| Open space inside park           | 1.0213 (1.0059, 1.0369) |
| Open space outside park          | 1.0166 (1.0066, 1.0327) |
| Forest inside park               | 0.9441 (0.9284, 0.9601) |
| Forest outside park              | 0.9167 (0.9032, 0.9304) |

Note: IRR = incidence rate ratios, CI = confidence interval.
4.1.3. Why does forest outside park have a stronger negative association with infection rate than forest inside park?

We found forest outside park to have a greater negative association with the infection rate than the forest inside park. One possible reason is that naturalistic green spaces are more likely to be found in forest outside park. Previous studies have suggested that naturalistic green spaces are more strongly associated with health benefits than urban green spaces (Allard-Poesi et al., 2022).

Second, US residents have a much higher level of exposure to forest outside park than forest inside park. According to our calculation in this study, population-weighted exposure to forest outside park is 10 times greater than exposure to forest inside park within 4 km (Fig. 9). The dominating supply of forest outside park might make the forest outside park have a greater impact than forest inside park.

Third, forest inside park, as a public space, might have a greater responsibility to accommodate more visitors and social gatherings than forest outside park, which makes users of forest inside park have a greater risk of infection than users of forest outside park. Mobility studies found that park visits were associated with higher SARS-CoV-2 infection rate and suggested that parks may serve as locations of virus transmission (DePhillipo et al., 2021; Praharaj & Han, 2021). Other studies suggest outdoor gatherings increase risk of respiratory disease outbreaks (Dixon et al., 2013; Domenech-Montoliu et al., 2021). Thus, increased risk social gatherings may counteract the other benefits of forest inside park.

Finally, the effect of forest inside park might be impacted by shut-down policies (Smith et al., 2021; Volenec et al., 2021). Many parks were fully or partly shut down during the pandemic (Smith et al., 2021).

4.1.4. Why do open spaces have a significantly positive and a non-significant association with SARS-CoV-2 infection rate?

We found open space inside park to be significantly associated with higher SARS-CoV-2 infection rate, and open space outside park to be non-significantly associated with infection rate. In this study, open space is defined as “… mostly vegetation in the form of lawn grasses. Impervious surfaces account for <20 % of total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in settings for recreation, erosion control, or aesthetic purposes” (NLCD, 2016). Thus, most open space, especially those in parks, are important places for recreational and social activities.

At the first glance, these findings are surprising given that many studies have found that open spaces are beneficial for health although their effects might be smaller than those of forest (please refer to 4.1.2). The most frequently reported benefits include reducing stress, fatigue, and negative emotions (Jiang, Schmillen, et al., 2019; Ulrich et al., 1991); promoting physical activities (Giles-Corti et al., 2005); promoting social cohesion (Jennings & Bamkole, 2019; Schmidt et al., 2019); and reducing the incidence of noninfectious chronic diseases, which mainly include cardiovascular diseases, stroke, cancers, and diabetes (Kondo et al., 2018). Open spaces, similar to other green spaces, might enhance human’s resistance to infection risk through enhancing mental health and then immunization functioning, and reduce solid aerosols (e.g., PM 2.5) in the air that might serve as a SARS-CoV-2 carrier as we discussed in 4.4.1.

However, our findings also echo many studies that reported open spaces have mixed impacts of open spaces on SARS-CoV-2 infection rate (Johnson et al., 2020; Klompmaker et al., 2021). It is possible that...
positive effects of open spaces have been significantly offset by their detrimental impacts. This interpretation is supported by studies that suggest outdoor social activities, such as mass gathering, walking, or partying, can lead to higher risk of infection (Doménech-Montoliu et al., 2021; Peng et al., 2022). Although open spaces can provide a relatively larger social distancing than indoor spaces, limited supply of open spaces in urban areas may make it hard for people to maintain safe social distancing all the time (Nobajas et al., 2020; Shoari et al., 2020). As the pandemic lasted longer than people expected, the pandemic fatigue may also influence people’s compliance of safe social distancing when they gathered in open spaces (Franzen & Wöhner, 2021; Shearston et al., 2021).

### 4.1.5. How to interpret forest-infection rate associations across five levels of urbanicity?

For the forest-infection rate associations across five levels of urbanicity, we found a key pattern: both forests variables have non-significant positive effects of open spaces have been significantly offset by their covariates at varying buffer distances, including state as a random effect (Model 4).

### 4.1.6. How to interpret forest-infection rate associations across five time periods?

For the associations across five time periods, we also found a key pattern: the negative associations between two forest variables and infection rate are stronger in an early time periods (period 2) but the negative associations largely remain significant in later time periods (period 3, 4, and 5).

One possible reason might be the early in the pandemic, the total number of infected people was small, and the virus was largely spread via social gatherings in private or institutional places (Leclerc et al., 2020; Thakar, 2020). Therefore, the total number of infected people were relatively small in green spaces, and it was easier for people to have recreational activities in green spaces while maintaining safe social distancing. In later stages, the pandemic had widely spread. The risk of infection largely increased due to a higher proportion of infected population and harder to keep safe social distancing.

One additional reason might be the pandemic fatigue developed overtime (Haktanir et al., 2021). The fatigue was mainly due to the increasing levels of mental and physical exhaustion overtime. Recent studies found decline in protective behaviors, lower perceived severity of SARS-CoV-2, and increase in the visit of retail and recreation locations compared to the early stage of the pandemic (Franzen & Wöhner, 2021; Haktanir et al., 2021; Shearston et al., 2021; Petherick et al., 2021; Maclntyre et al., 2021).

Nevertheless, we should emphasize again that the negative associations remain significant across five periods, which suggest impacts of forest on relieving infection rate can keep being robust through a long period of time.

### 4.1.7. Forrest within walking distance associated with reduce infection rate

In this study, we identified an optimal distance to forested areas by considering the spatial distribution of populations and forests across counties. We found that negative associations between forests and infection rate are strongest for forests within moderate walking distance (forest outside park ≤ 1,200 m and forest inside park ≤ 600 m). The radius of optimal buffer zones largely matches the most favorable walking distances in the United States (Yang & Diez-Roux, 2012). Our findings are consistent with other studies that report the significant health benefits of nearby forest and other green spaces (Corraliza & Collado, 2011; Cox et al., 2017; Lee et al., 2019; Oh et al., 2017).

The reason why nearby forest has a stronger effect of relieving infection risk can be interpreted as direct and indirect reasons. For the direct reason, nearby forest is more frequently visited by residents than distant forest. More frequent visits can lead to better mental and physical health and then better immunization functioning to resist the infection risk (Kuo, 2015; Roviolesi et al., 2021). Past studies suggest the frequency of green space visit declines with increasing distance (Coombes et al., 2010; Zender & Ward Thompson, 2017). An increase in people walking to nearby green space during SARS-CoV-2 pandemic was also observed (Ugolini et al., 2020). For the second reason, nearby forest has a stronger effect on reducing concentration of air pollutants in residential areas, such as PM_{2.5} and PM_{10}, that might be carriers of SARS-CoV-2 virus (Czwojdzinska et al., 2021; Nor et al., 2021; Qu et al., 2019). Thus, nearby forest, no matter they are accessible or not by public, can reduce the infection risk of people who live in nearby neighborhoods.

### 4.2. Significance and contributions to knowledge and practice

To our best knowledge, this study is one of the first nationwide studies investigating the relationships between different types of green spaces and SARS-CoV-2 infection rate. The significance and contribution of this study mainly include the following aspects:

The control of key confounding factors in this study is more comprehensive and rigorous than in previous studies. We include socioeconomic and demographic factors, pre-existing chronic disease.

### Table 5

| Variables | Buffer size (m) | Beta | SE | z-value | p-value |
|-----------|----------------|------|----|---------|---------|
| Forest inside park | 100 | -0.014 | 0.010 | −1.400 | 0.161 |
| | 200 | -0.020 | 0.009 | −2.119 | 0.034* |
| | 400 | -0.022 | 0.009 | −2.348 | 0.012** |
| | 600 | -0.023 | 0.009 | −2.629 | 0.009** |
| | 800 | -0.023 | 0.009 | −2.648 | 0.008** |
| | 1,000 | -0.023 | 0.009 | −2.616 | 0.009** |
| | 1,200 | -0.022 | 0.009 | −2.610 | 0.009** |
| | 1,400 | -0.022 | 0.009 | −2.507 | 0.009** |
| | 1,600 | -0.022 | 0.009 | −2.597 | 0.009** |
| | 1,800 | -0.022 | 0.009 | −2.615 | 0.009** |
| | 2,000 | -0.023 | 0.009 | −2.627 | 0.009** |
| | 2,500 | -0.023 | 0.009 | −2.679 | 0.007** |
| | 3,000 | -0.024 | 0.009 | −2.758 | 0.006** |
| | 3,500 | -0.025 | 0.009 | −2.872 | 0.004** |
| | 4,000 | -0.026 | 0.009 | −2.954 | 0.003** |
| Forest outside park | 100 | -0.085 | 0.008 | −11.353 | <0.0001*** |
| | 200 | -0.087 | 0.008 | −11.244 | <0.0001*** |
| | 400 | -0.089 | 0.008 | −11.104 | <0.0001*** |
| | 600 | -0.090 | 0.009 | −11.510 | <0.0001*** |
| | 800 | -0.091 | 0.008 | −11.578 | <0.0001*** |
| | 1,000 | -0.091 | 0.008 | −11.635 | <0.0001*** |
| | 1,200 | -0.091 | 0.008 | −11.669 | <0.0001*** |
| | 1,400 | -0.091 | 0.008 | −11.692 | <0.0001*** |
| | 1,600 | -0.091 | 0.008 | −11.701 | <0.0001*** |
| | 1,800 | -0.091 | 0.008 | −11.701 | <0.0001*** |
| | 2,000 | -0.091 | 0.008 | −11.705 | <0.0001*** |
| | 2,500 | -0.090 | 0.008 | −11.704 | <0.0001*** |
| | 3,000 | -0.089 | 0.008 | −11.694 | <0.0001*** |
| | 3,500 | -0.089 | 0.008 | −11.677 | <0.0001*** |
| | 4,000 | -0.089 | 0.008 | −11.641 | <0.0001*** |
factors, political and policy factors, healthcare and testing factors, behavioral factors, and climate and environmental factors. In addition, control of bias caused by spatial autocorrelation can further enhance the validity of our findings. After controlling all these factors in statistical analysis, the association between green spaces and infection rate was found to be statistically independent and significant.

Numerous studies have found green space, as a general type of land cover, can be beneficial for health and society already recognized it. This type of generic finding is not informative enough to guide specific policymaking and planning interventions (Lu et al., 2021; Klompmaker et al., 2021). This study moves beyond this limitation by revealing the relationships between various types of green spaces and SARS-CoV-2 infection rate and identifying forest as the key type of green space. Further, the study reveals the relationship between forest and SARS-CoV-2 infection rate across five levels of urbanicity and five periods of time. Lastly, the study reveals the relationship between population-weighted forest and infection rate thus identifying optimal buffer zones that are associated with the lowest inflection risk. Through all these key steps, this study provides more specific evidence to guide practice to reduce the risk of airborne infectious diseases.

4.3. Limitations as future research opportunities

This study has several limitations that should be further investigated
by future studies. First, this study presents correlational, rather than causal, findings. A causal relationship between green spaces and infection rate is plausible, based on the mechanisms proposed and a wealth of previous research. Future researchers should conduct experimental studies, including laboratory or natural experiments, to confirm these causal relationships (Jiang et al., 2021; Tyrvainen et al., 2014).

Second, our study focuses on investigating and interpreting the potential effects of green spaces, especially forest, on the infection rate. We find, however, that many other factors are also significantly associated with infection rate, such as the Gini index, overcrowded housing, political factors, numbers of essential workers, modes of transport used for commuting, and public mask mandates. These factors all have significant potential to be the focal point of future studies.

Third, we did not use data collected in 2021, as vaccination programs were implemented in the early months of 2021, likely confounding the relationship between green spaces and infection rate (BBC, 2020). Here again, we see this as an important opportunity for future research. It is necessary to understand the extent to which the vaccination rate alters the relationship between green spaces and the SARS-CoV-2 infection rate.

Fourth, some variables are time-sensitive, e.g., testing rate and social distancing policies. The temporal dynamics were considered by splitting the data into five time periods in this study. Future research may use longitudinal data analysis (e.g., multilevel model) to model such temporal dynamics.

Finally, although this study employed an ecological design that includes a series of population-weighted assessments, inferences cannot be made about individual levels of infection based on aggregate data gathered at the county level. This challenge is difficult to overcome, given the difficulty of acquiring personal SARS-CoV-2 data. Perhaps researchers in countries where individual data is available can address this limitation.

5. Conclusion

This one-year study is one of a few nationwide studies investigating the relationships between different types of green spaces and the SARS-CoV-2 infection rate. The consideration of spatial autocorrelation, population-weighted measure, and control of a variety of covariates adds to the study’s validity. As the whole world continue to battle the SARS-CoV-2 pandemic and prepare preventive solutions for future health crises, we urge them to prioritize equitable and accessible green spaces, especially those that contain forest, as a critical public health strategy.

Funding

This work was supported by the University of Hong Kong [grant numbers 102010054.088616.01100.302.01] and [grant numbers 104006587.088616.01100.301.01] and the Research Grants Council of the Hong Kong SAR [grant number CityU11207520].

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.

Data availability

Data will be made available on request.

Acknowledgments

We are grateful that this study received enormous trust and support from the following leaders and colleagues at the University of Hong Kong: Dean Chris Webster, Associate Dean Mathew Pryor, Associate Dean Wilson W.S. Lu of the Faculty of Architecture, Head Eric H. Schuldenfrei of the Department of Architecture, Head Ivan Valin of the Division of Landscape Architecture. We appreciate valuable comments and suggestions from Chao Song, Matthew Browning, Lan Luo, and Wenyan Xu.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2022.104583.

References

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., ... Zand, R. (2020). Racial, economic, and health inequality and COVID-19 infection in the United States. Journal of Racial and Ethnic Health Disparities, 8, 1–11.
