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.

Cities 118 (2021) 103396

Contents lists available at ScienceDirect

Cities

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

What county-level factors influence COVID-19 incidence in the United States? Findings from the first wave of the pandemic

Lan Wang [a,*], Surong Zhang [a], Zilin Yang [b], Ziyu Zhao [c], Anne Vernez Moudon [d], Huasen Feng [a,e], Junhao Liang [a,e], Wenyao Sun [a], Buyang Cao [a,e,*]

[a] College of Architecture and Urban Planning, Tongji University, China
[b] Cultural Heritage Studies Institute of Archaeology, University College London, United Kingdom of Great Britain and Northern Ireland
[c] Department of City and Regional Planning, Cornell University, United States of America
[d] Urban Form Lab, Department of Urban Design and Planning, University of Washington, United States of America
[e] College of Software Engineering, Tongji University, China

ARTICLE INFO

Keywords:
COVID-19
Influential factors
Lockdown
Communicable disease prevention
Structural equation modeling (SEM)

ABSTRACT

Effective control of the COVID-19 pandemic via appropriate management of the built environment is an urgent issue. This study develops a research framework to explore the relationship between COVID-19 incidence and influential factors related to protection of vulnerable populations, intervention in transmission pathways, and provision of healthcare resources. Relevant data for regression analysis and structural equation modeling is collected during the first wave of the pandemic in the United States, from counties with over 100 confirmed cases. In addition to confirming certain factors found in the existing literature, we uncover six new factors significantly associated with COVID-19 incidence. Furthermore, incidence during the lockdown is found to significantly affect incidence after the reopening, highlighting that timely quarantining and treating of patients is essential to avoid the snowballing transmission over time. These findings suggest ways to mitigate the negative effects of subsequent waves of the pandemic, such as special attention of infection prevention in neighborhoods with unsanitary and overcrowded housing, minimization of social activities organized by neighborhood associations, and contactless home delivery service of healthy food. Also worth noting is the need to provide support to people less capable of complying with the stay-at-home order because of their occupations or socio-economic disadvantage.

1. Introduction

Since December 2019, a novel beta-coronavirus, known as the severe acute respiratory syndrome coronavirus, has spread throughout the world, bringing countries to a standstill and dragging the global economy into a dramatic recession (American Journal of Managed Care, 2020). The disastrous COVID-19 outbreak has been reviewed and studied by scholars from different disciplines, such as epidemiology (Chen et al., 2020; Millett et al., 2020; Pan et al., 2020), biology (Wang et al., 2020; Wu & McGoogan, 2020), medicine (Cai, 2020; Sinclair & Abdelhafiz, 2020; Vardavas & Nikitara, 2020; Whittle & Diaz-Artiles, 2020), demography (Borjas, 2020; Dowd et al., 2020; Pareek et al., 2020), and psychiatry (Yang et al., 2020). These existing studies have identified certain factors significantly correlated with COVID-19 infections, including public health interventions, age, gender, minority groups, comorbidity and smoking etc. They could basically be categorized into three aspects under a research framework for communicable disease prevention: protection of vulnerable populations, intervention in transmission pathways, and provision of healthcare resources. Nevertheless, the impact of more variables in these three aspects, such as housing conditions, home-based occupations, community activities and the number of Primary Care Physicians, have not been systematically studied via an integrated model.

To unravel the impact of a complexity of potential factors on COVID-19 incidence and its working mechanism, this paper includes more variables under the research framework for a systematic analysis. Counties with more than 100 confirmed cases in the United States are selected as the spatial analytical unit. We collected datasets for modeling analysis using two scenarios, one during the April 2020 lockdown, and one following the July 2020 reopening, a period roughly covering the

* Corresponding authors at: College of Architecture and Urban Planning, Tongji University, 1239 Siping Road, Shanghai 200082, China.
E-mail addresses: wanglan@tongji.edu.cn (L. Wang), caobuyang@tongji.edu.cn (B. Cao).

https://doi.org/10.1016/j.cities.2021.103396
Received 7 August 2020; Received in revised form 9 June 2021; Accepted 21 July 2021
Available online 28 July 2021
0264-2751/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

first wave of the COVID-19 pandemic in the United States (Ioannidis et al., 2021; YouTube, 2020). The novelty of this paper lies in two aspects. On the one hand, we have developed a new research framework that uncovers several new variables, the significance of which has not been widely studied in the existing literature. We have confirmed their significant association with COVID-19 infections by linear regression models and structural equation modeling (SEM). On the other hand, based on the datasets from two different days, we have further analysed the influence pathways among the variables and identified the factors affecting COVID-19 transmission during both the lockdown and reopening periods by comparing the differences between the two scenarios.

The paper first reviews scholarship on factors associated with COVID-19 infections, and then develops a research framework for modeling, which is illustrated in the third section. The linear regression models and SEM are developed and analysed in the fourth and fifth sections. Research findings are then compared with existing studies and relevant suggestions are provided for policy-making on prevention and control of the COVID 19 pandemic.

2. Literature review

Scholars have indicated how differences in demographic features, socio-economic attributes, and individual health conditions significantly affect COVID-19 incidence or mortality. Ages under 18 or over 65, as shown in research on clinical characteristics of infected cases, are proven to be risk factors (Sinclair & Abdelhafiz, 2020; Whittle & Diaz-Artiles, 2020). Older patients might face greater risk of developing a critical or mortal condition (Novel, 2020; Wang et al., 2020; Wu et al., 2020; Zheng et al., 2020; Zhou et al., 2020). Regarding gender, men are at higher risk of COVID-19 infection and mortality than women (Cai, 2020; Dowd et al., 2020; Pareek et al., 2020; Whittle & Diaz-Artiles, 2020; Zheng et al., 2020). In relation to ethnicity, analysing confirmed cases of New York City neighborhoods, Borjas (2020) discovers that probability of infection is higher in neighborhoods with a large population of African Americans or immigrants. Millett et al. (2020) also confirms the vulnerability of African Americans by assessing impacts of the virus on different communities via regression analysis. In terms of income, studies have found that neighborhoods with low-income households present higher COVID-19 incidence rates in counties of the U.S. (Mollalo et al., 2020; Whittle & Diaz-Artiles, 2020). Meanwhile, several comorbidities, including hypertension, diabetes, cardiovascular and respiratory diseases, could greatly affect the prognosis of COVID-19 (Fang et al., 2020; Guo et al., 2020; Sinclair & Abdelhafiz, 2020; Wang et al., 2020; Zheng et al., 2020). Smoking and obesity are also found to be associated with the prevalence and adverse outcomes of COVID-19 (Cai, 2020; Kass et al., 2020; Lighter et al., 2020; Vardavas & Nikitara, 2020; Zheng et al., 2020). Other risk factors may include air pollution. Air pollutants are risk factors for respiratory infections because they carry microorganisms and affect human immunity (Becker & Soukup, 1999; Xu et al., 2016). COVID-19 is a respiratory disease, and significantly positive associations, including a higher mortality risk, have been observed between air pollution and the increase in COVID-19 confirmed cases (Travaglio et al., 2021; Xu et al., 2020; Zhang et al., 2020; Zhu et al., 2020). Although several studies have found that reducing air pollution may contribute to control of the pandemic and increase the coping capacity of citizens, this important finding is not yet well-explored and warrants more evidence (Sharifi & Khavarian-Garmsir, 2020).

A variety of studies focus on factors related to population density, social activities and healthcare resources. Stier et al. (2020) measure the relationship between city size and the COVID-19 attack rate in 201 U.S. cities and find that populous urban areas have higher infection rates. Borjas (2020) also discovers that neighborhoods with a high population density and large household size are more susceptible to COVID-19 infection. Social activities, such as working, commuting, shopping and public gathering, may lead to cluster transmission of COVID-19 (Lau et al., 2020; Qi et al., 2020). Mobility in these social activities is highly correlated with the number of COVID-19 cases and the established inter-provincial migrant networks in China played an important role in the spread of COVID-19 (Qi et al., 2020; Shen, 2020). Badr et al. (2020) demonstrate the positive correlation between mobility patterns and COVID-19 case growth in the U.S. counties by calculating the Pearson correlation coefficient between these two variables. Thus, travel restrictions and social distancing policies, which regulate human mobility, are helpful for controlling the spread of COVID 19 (Kraemer et al., 2020). Adequate provision of healthcare resources, including the number of licensed or ICU beds and the number of nurse practitioners, is found to be negatively associated with COVID-19 incidence (Grasselli et al., 2020; Li et al., 2020; Pan et al., 2020).

The above factors associated with COVID-19 incidence can be categorized into three aspects under a research framework for communicable disease prevention: protection of vulnerable populations, intervention in transmission pathways, and provision of healthcare resources. Nevertheless, we have identified additional factors in these three aspects that might be related to pandemic transmission but which have not yet been explored in depth in the existing literature. Although some studies show that COVID-19 could be related to a rise in unemployment rates (Coibion et al., 2020; Kawohl & Nordt, 2020), few analyse unemployment rates as one of the explanatory variables of COVID-19 infection. Unemployed people may have lower incomes and lack access to health protection; however, they may stay at home most of the time during the pandemic. Thus, it is difficult to determine whether unemployed people are more vulnerable to COVID-19, which requires further examination. In addition, studies on the inter-relationship between occupation and risk of virus exposure show that people working from home are less exposed to pathogens (Baker, 2020; Dingel & Neiman, 2020), and a lower risk of virus exposure is thought to be associated with fewer disease infections (Smith et al., 2005). The proportion of home-based occupations may, therefore, be a potential predictor of COVID-19 incidence but, to our knowledge, no literature has explored this relationship.

Factors related to community conditions, including the housing environment and neighborhood associations, are influential factors for the transmission of infectious diseases (WHO, 2018). Individuals in crowded dwellings are more susceptible to infectious diseases (Baker et al., 2013). Meanwhile, neighborhood associations are assumed to pass on information and improve adherence to transmission-reduction policies (Whittle & Diaz-Artiles, 2020), although these associations may hold community gathering events leading to person-to-person contact. Therefore, the effect of neighborhood association rates on COVID-19 transmission remains unpredictable. However, the two variables receive little attention in relation to their roles in COVID-19 infections.

It should be noted that the food environment may also influence the transmission and infection of COVID 19. In neighborhoods with ready access to stores selling fresh and healthy food, people tend to have healthier diets, strengthening the immune system (Babey et al., 2008; Glanz et al., 2005; Morland & Evenson, 2009) and reducing rates of comorbidity (Leddy et al., 2020). This may be associated with lower numbers of COVID 19 infections. On the other hand, shopping in crowded supermarkets to buy fresh and healthy food may increase the chance of close contact, thus increasing the risk of virus transmission. Whether the food environment will have a positive or negative effect on the spread of COVID-19 is another question waiting to be explored.

Additionally, the number of Primary Care Physicians is another factor that has not received much attention, and we believe it should also be a potential indicator of COVID-19 incidence, as Primary Care Physicians play an important role in general public healthcare, especially in relation to comorbidity (Starfield et al., 2003). They are usually the medical professionals sought by patients when they first encounter health problems, and provide initial diagnosis and direct care for patients with different types of diseases.

While omitting consideration of the aforementioned potential factors affecting COVID 19 infections, existing studies also have limitations in their research design and methodology, since most are cross-sectional rather than longitudinal studies. Considering the rapidly evolving nature of COVID 19, more longitudinal or time-series studies should be conducted. Further, current literature generally applies regression models and spatial statistical models for correlation analysis (Mollalo et al., 2020; Kass et al., 2020; Luo et al., 2020), whereas only a few explore the causal relationship between explanatory variables and COVID-19 infection rate and mortality rate via SEM (Hamidi et al., 2020). Therefore, selecting two dates representing the lockdown and reopening periods respectively, this study investigates aforementioned factors that are potentially associated with COVID-19 transmission via linear regression models, and explore causal pathways between variables in COVID 19 infections in different periods via SEM.

3. Research design and methods

3.1. Research framework

As discussed above, the influential factors of COVID-19 incidence identified in the existing literature can be categorized into three aspects: protection of vulnerable populations, intervention in transmission pathways, and provision of healthcare resources (see Fig. 1).

Additional predictors relevant to these three aspects and worth exploring are included to unravel their association with COVID-19 incidence, namely, unemployment, housing condition, neighborhood associations, home-based occupation, food environment and Primary Care Physicians. The outcome variables are COVID-19 incidence from different scenarios ('stay-at-home order' in place vs. lifted) calculated by dividing the cumulative number of new COVID-19 cases by the total population in U.S. counties. We hypothesize that variables significantly affecting COVID-19 incidence vary across different scenarios during the lockdown and after the reopening.

3.2. Case selection and data collection

We adopt counties in the U.S. as the analytical spatial units because this is the smallest geographical unit for which consistent COVID-19 data can be obtained. Aggregation bias is less likely to occur than in analyses performed on a larger spatial unit.

3.2.1. Outcome variables

We have chosen incidence as of April 27th, 2020 and incidence as of July 8th, 2020 as outcome variables representing scenarios during the lockdown and following the reopening, respectively. These two days are chosen based on the timing of issuing relevant policy and consider the delay between policy announcement and its implementation. April 27th is six weeks after the social distancing guidelines were issued and two weeks after signing of the Presidential Disaster Declaration for all 50 states. July 8th is about six weeks after all 50 states lifted their restrictions. Data for the outcome variables is obtained from the Johns Hopkins University GitHub open-source project (2020).

Moreover, all counties involved in the analysis have more than 100 confirmed cumulative cases as of April 27th, 2020. This indicates the massive transmission in these counties, which could provide suitable samples for exploring influential factors of transmission. A total of 634 counties are selected and illustrated in Fig. 2.

3.2.2. Explanatory variables

County-level data for the explanatory variables is categorized according to the research framework presented in Section 3.1. The data relating to vulnerable populations are obtained from a variety of data sources:

- Data on gender, age, and ethnicity are collected from the website of County Health Ranking, provided by Census Population Estimates (2016).
- Data on income ratio, unemployment rate, and 20th percentile income are obtained from the website of County Health Ranking, hosted by the Bureau of Labor Statistics (2018).
- Data on adult smoking rates are acquired from the website of County Health Ranking, provided by the Behavioral Risk Factor Surveillance System (2017).
- Data on adult obese and diabetic rates information are obtained from the website of County Health Ranking, hosted by the United States Diabetes Surveillance System (2016).

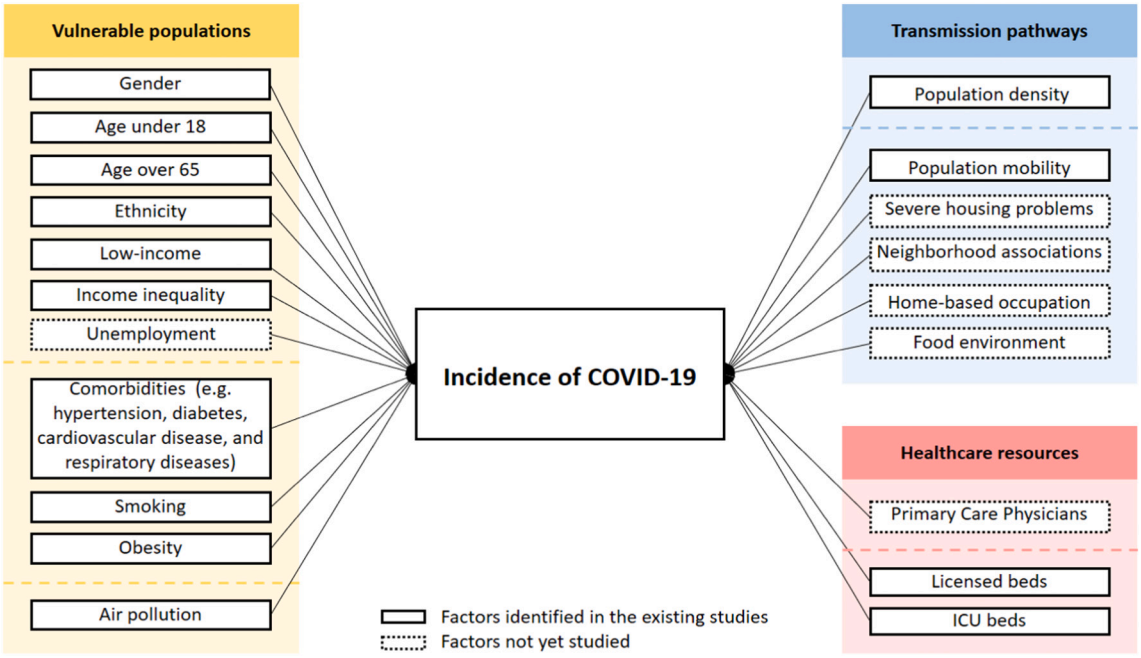

Fig. 1. Research framework of influential factors of COVID-19 incidence.

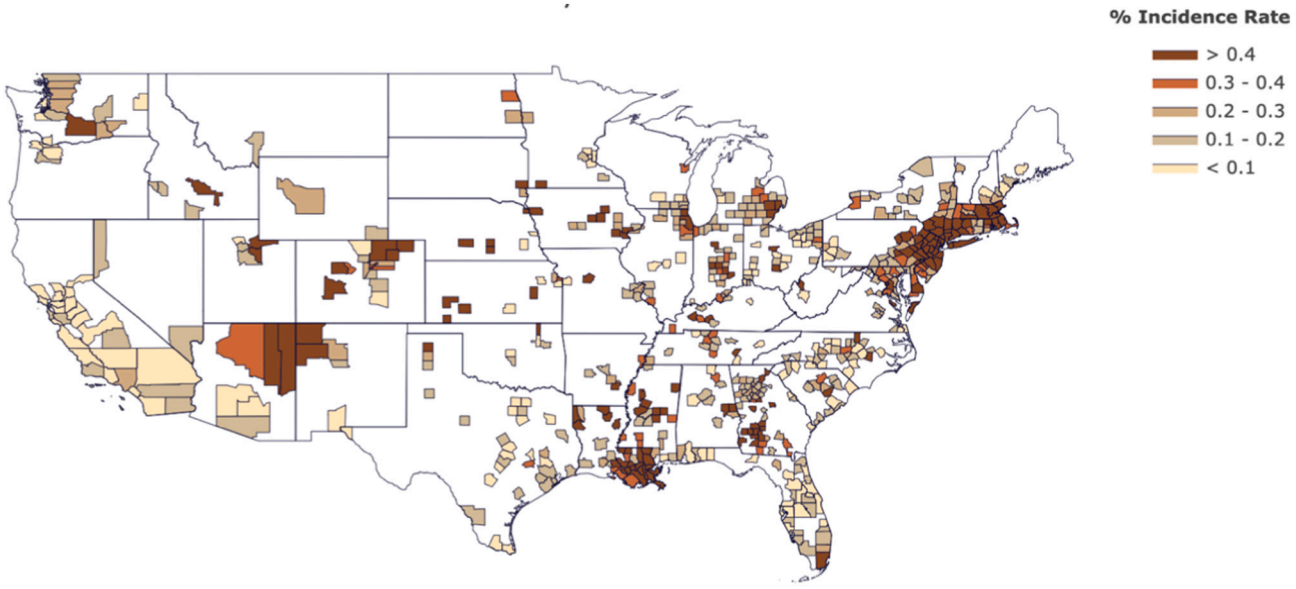

Fig. 2. Incidence rate in counties in the U.S. as of April 27th, 2020.

Table 1
The definition and descriptive statistics of the explanatory variables (full sample, $n = 643$).

	Variables	Definition	Unit	Mean	Std. dev	Min	Max
Vulnerable population	% Female	The percentage of the population that is female	NA	50.850	1.553	34.160	53.850
	% Under 18	The percentage of the population that is under 18	NA	23.010	3.618	7.070	34.190
	% Over 65	The percentage of the population that is over 65	NA	15.110	3.939	7.420	56.310
	% African American	The percentage of the population that is African American	NA	8.980	15.031	0.210	82.250
	% Asian	The percentage of the population that is Asian	NA	2.160	4.376	0.220	42.900
	20th percentile income	20th percentile of median household income	Dollars	22,741.000	7512.513	7601.000	60,890.000
	Income ratio	Ratio of household income at the 80th percentile to income at the 20th percentile	NA	4.500	0.740	3.220	8.930
	% Unemployed	The percentage of population aged 16+ unemployed and looking for work	NA	4.800	1.566	2.030	23.520
	% Diabetic	The percentage of adults aged 20 and above with diagnosed diabetes	NA	11.000	2.37	3.000	17.000
	% Smokers	The percentage of adults that reported currently smoking	NA	17.000	3.422	7.000	27.000
	% Obese	The percentage of adults that report BMI $\geq$ 30	NA	30.000	5.046	13.000	42.000
	Average daily PM2.5	Average daily amount of fine particulate matter in micrograms per cubic meter	Per cubic meter	9.800	1.617	0.000	15.400
Transmission pathways	Population density	Measurement of population per unit area	Per square mile	309.844	3192.247	4.364	71,510.017
	Mobility index	the number of people not staying home divided by the number of people staying home	NA	2.650	0.644	0.771	6.956
	% Severe housing problems	The percentage of households with at least 1 of 4 housing problems: overcrowding, high housing costs, or lack of kitchen or plumbing facilities	NA	16.000	4.369	8.000	35.000
	Association rate	The number of neighborhood associations per 10,000 population	NA	9.830	3.363	1.890	27.490
	% Education services	The percentage of the total employment that is Education Services.	NA	0.017	0.015	0.000	0.116
	% Finance and insurance	The percentage of the total employment that is Finance and Insurance	NA	0.037	0.022	0.000	0.256
	% Information	The percentage of the total employment that is Information	NA	0.010	0.009	0.000	0.097
	% Professional, scientific, and technical services	The percentage of the total employment that is Professional, Scientific, and Technical Services	NA	0.049	0.033	0.000	0.243
	% Management of companies and enterprises	The percentage of the total employment that is Management of Companies and Enterprises	NA	0.006	0.010	0.000	0.083
	Food environment index	Indicator of access to healthy foods - 0 is worst, 10 is best	NA	7.800	1.119	0	10
Healthcare resources	PCP ratio	Population divided by the number of Primary Care Physicians	NA	1484.000	1519.317	0.000	20,781.000
	Licensed beds rate	The number of licensed beds divided by population	per capita	0.003	0.003	0.000	0.032
	ICU beds rate	The total number of ICU beds, psychiatric ICU beds, and Detox ICU beds divided by population	Per capita	0.000	0.000	0.000	0.003

- Data on the average daily PM2.5 are obtained from the website of County Health Ranking, hosted by Environmental Public Health Tracking Network (2014).

Data for the variables depicting the transmission pathways are mainly related to the built environment and socio-economic activity, which are obtained from the following sources:

- The population density data are downloaded from the website of the Open Data Network, provided by the U.S. Census American Community Survey (2018).
- Data on the percentage of households with severe housing problems are acquired from the website of County Health Ranking, provided by the Comprehensive Housing Affordability Strategy (CHAS) (2016).
- Data on the Food Environment Index are collected from the website of County Health Ranking, provided by USDA Food Environment Atlas (2015).
- Data on neighborhood association are collected from the website of County Health Ranking, provided by County Business Patterns (2017).
- Data on occupations are obtained from the website of the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce (2020). Here we use five occupations that are more amenable to home-based work, including: 1) education services, 2) finance and insurance, 3) management of companies and enterprises, 4) information, and 5) professional, scientific, and technical services.
- Trip by distance data provided by U.S. Department of Transportation (2020) are used to compute the Mobility Index defined by the number of people in a county not staying home divided by the number of people staying home.

For healthcare resources, the dataset on the ratio of population to Primary Care Physicians, ICU beds, and licensed beds is gathered from the website of ESRI COVID-19 Overview (2020), released on March 21st, 2020.

The definition and descriptive statistics of explanatory variables are presented in Table 1.

3.3. Analytical methods

To fully test our hypotheses, the analytical approach is divided into two stages. In the first stage, all the explanatory variables are tested for significance as predictors of incidence as of April 27th and incidence as of July 8th, respectively. We apply the linear regression models to disclose correlations between the explanatory variables and COVID-19 incidence. SPSS 24 statistical software is used to conduct the model fitting and data analyses. A Pearson correlation analysis is employed to ensure the absence of high correlation between the explanatory variables. All variables are normalised before being fed to the model.

In the second stage, only the explanatory variables that have statistically significant relationships to the outcome variables in the linear regression models are included in the SEM, which could effectively verify the causal link in our hypothesis. The SEM in this study is estimated using Amos 24. Since COVID-19 incidence as the dependent variable did not conform to a normal distribution, we chose generalized least squares for the estimation.

4. Results overview

4.1. Results from the linear regression modeling

The following overviews are based on the results presented in Table 2. We fit the linear regression models using the COVID-19 incidence data of counties ($N = 643$) collected on April 27th, 2020 (model 1) and July 8th, 2020 (model 2). The R square of model 1 is 0.543, indicating that a large percentage of the variance in COVID-19 incidence is explained by the explanatory variables included in the analysis. The R square of model 2 is 0.526, also indicating good fitting effect. This demonstrates that the proposed models can reveal the characteristics of COVID-19 infection in the United States.

4.1.1. Insights into vulnerable population

Vulnerability to COVID-19 infection varies among population with different demographic and socio-economic characteristics. Both models reveal that females are generally at lower risk of COVID-19 infection. People under 18 or over 65 are more vulnerable to COVID-19 infection. In relation to ethnicity, the percentage of African Americans is found to be significantly positively related to COVID-19 incidence, while the

Table 2
Results from the linear regression analysis between the county-level explanatory variables and COVID-19 incidence as of April 27th and July 8th (Full sample, $n = 643$).

	Variables	Model 1: April 27th			Model 2: July 8th		
		B	t	P	B	t	P
	(Constant)	−0.008	−2.130	0.034*	−0.009	−1.629	0.104
Vulnerable population	% Female	−0.022	−8.474	0.000***	−0.036	−8.399	0.000***
	% Under 18	0.010	4.384	0.000***	0.034	9.190	0.000***
	% Over 65	0.018	5.484	0.000***	0.024	4.496	0.000***
	% African American	0.005	3.597	0.000***	0.006	2.701	0.007**
	% Asian	−0.007	−2.866	0.004**	−0.009	−2.295	0.022**
	20th percentile income	0.017	5.936	0.000***	0.013	2.979	0.003**
	Income ratio	0.014	5.904	0.000***	0.018	4.630	0.000***
	% Unemployed	−0.012	−3.791	0.000***			
	% Smokers	0.005	2.574	0.010**			
	Average daily PM2.5				−0.011	−3.733	0.000***
Transmission pathways	Population density	0.085	18.659	0.000***	0.117	15.510	0.000***
	Mobility index				0.010	2.737	0.006**
	% Severe housing problems	0.009	5.488	0.000***	0.016	5.856	0.000***
	Association rate	0.007	4.116	0.000***	0.018	6.633	0.000***
	% Finance and insurance				0.010	2.437	0.015**
	% Professional, scientific, and technical services	−0.011	−5.047	0.000***	−0.011	−3.090	0.002**
	Food Environment index	0.008	3.313	0.001**	0.009	2.316	0.021**
Healthcare resources	PCP ratio	0.017	6.079	0.000***	0.034	7.322	0.000***

* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

percentage of Asians is found to be significantly negatively related to COVID-19 incidence. Income (using 20th Percentile Income as the proxy) and income inequality (using income ratio as the proxy) show a significantly positive association with COVID-19 incidence in both models. The percentage of unemployed people and the percentage of adult smokers are only significant in model 1, while average daily PM2.5 is only significant in model 2.

4.1.2. Formation of transmission pathways

Interestingly, Mobility Index and the percentage of employment in Finance and Insurance are not significant in model 1, but found to be significantly positively correlated with COVID-19 incidence in model 2. The results of other variables in terms of transmission pathways are consistent in both two models. We find that population density is a very strong influential factor of COVID-19 infection with the largest standard coefficients. The higher the population density, the higher the infectious probability will be. The percentage of households with severe housing problems, the neighborhood association rate and the Food Environment Index are found additionally to correlate positively and significantly with COVID-19 incidence. The percentage of employment in Professional, Scientific and Technical Services is negatively correlated with COVID-19 incidence.

4.1.3. Impacts of healthcare resources

The ratio of population to Primary Care Physicians is positively associated with COVID-19 incidence in both models; i.e., the higher the number of Primary Care Physicians, the lower the number of COVID-19 cases.

4.2. Results from the structural equation modeling (SEM)

Based on results from the linear regression models, the explanatory variables found to be significantly associated with COVID-19 incidence are included in the SEM to disclose causal relationships between the variables. Fig. 3 shows the best fit model from our SEM analysis, which has an acceptable low chi-square relative to the degrees of freedom (chi-square/degrees of freedom = 2.322) and an insignificant p value (p value = 0.128 > 0.05), indicating a good model fit. Table 3 presents the direct and indirect effects of county-level explanatory variables on COVID-19 incidence for both April 27th and July 8th.

Our findings indicate that all the variables originally significantly associated with COVID-19 incidence in the linear regression model for April 27th have the significant direct effects on COVID-19 incidence as of April 27th in the SEM, and the positive and negative direction of the effect remains consistent with the results from the regression model. Of all the variables, population density has the largest direct effect on incidence as of April 27th; its factor loading being the only one above 0.5.

The results of the SEM also reveal some interesting findings that have not been fully explored in the linear regression model. The incidence as of April 27th is found to have a significant and very large direct effect on the incidence as of July 8th; its factor loading far outweighing the other variables. Notably, seven variables (i.e., the percentage of African American, the percentage of Asian, 20th Percentile Income, Income Ratio, the percentage of unemployed people, the percentage of employment in Professional, Scientific, and Technical Services, and Food Environment Index) only have significant direct effect on the incidence as of April 27th, while three variables (i.e., average daily PM2.5, Mobility Index and the percentage of employment in Finance and Insurance) only have significant direct effect on the incidence as of

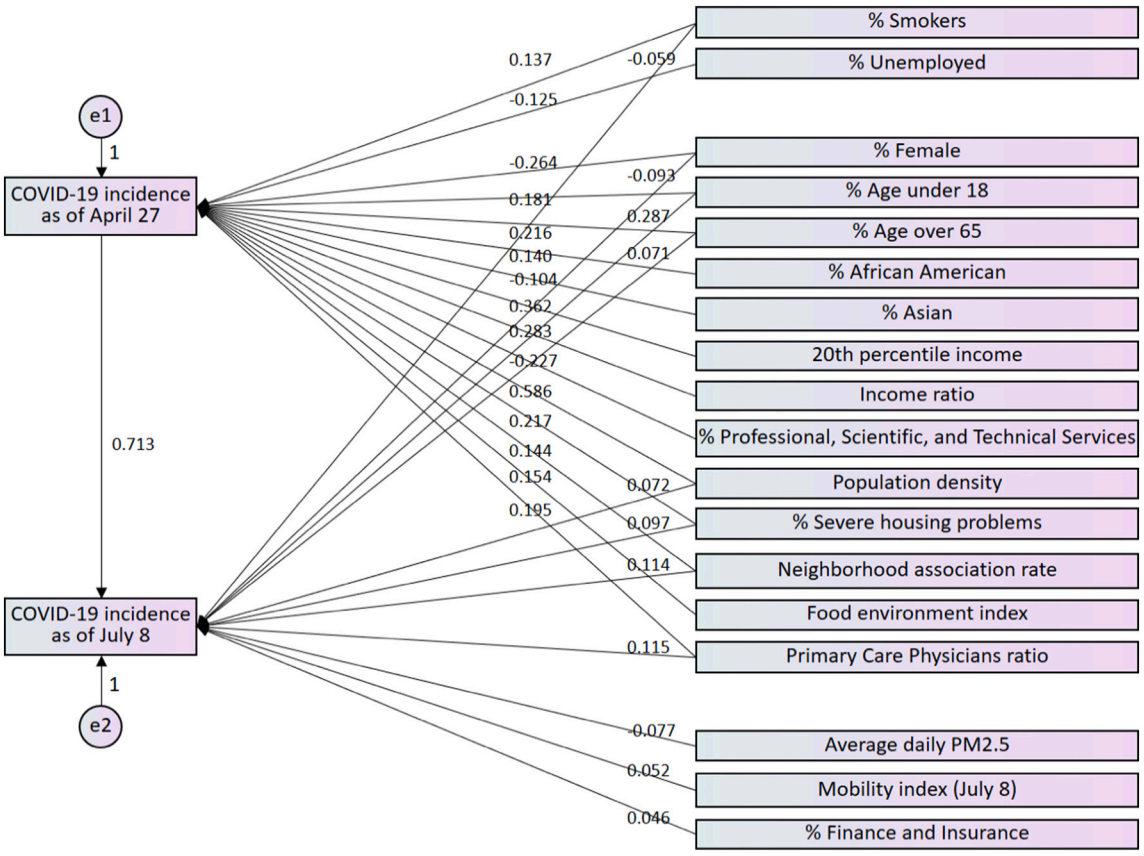

Fig. 3. Causal path diagram for COVID-19 incidence as of April 27th and July 8th in terms of the county-level explanatory variables (correlational arrows among explanatory variables are not shown to ensure simplicity and clarity).

Table 3
Direct and indirect effects of the county-level explanatory variables on COVID-19 incidence as of April 27th and July 8th (full sample, $n = 643$).

	Variables	Direct effect on incidence as of April 27th (i.e., indirect effect on incidence as of July 8th)		Direct effect on incidence as of July 8th	
		Estimate	p value	Estimate	p value
Vulnerable population	% Female	−0.264	***	−0.093	***
	% Under 18	0.181	***	0.287	***
	% Over 65	0.216	***	0.071	**
	% African American	0.140	***	0.019	0.495
	% Asian	−0.104	**	−0.011	0.672
	20th percentile income	0.362	***	−0.068	0.115
	Income ratio	0.283	***	0.056	0.114
	% Unemployed	−0.125	***	0.010	0.710
	% Smokers	0.137	**	−0.059	*
	Average daily PM2.5	−0.041	0.178	−0.077	***
Transmission pathways	Population density	0.586	***	0.072	**
	Mobility index			0.052	*
	% Severe housing problems	0.217	***	0.097	**
	Association rate	0.144	***	0.114	***
	% Professional, scientific, and technical services	−0.227	***	0.009	0.772
	% Finance and insurance	0.023	0.478	0.046	**
	Food environment index	0.154	***	−0.017	0.588
Healthcare resources	PCP ratio	0.195	***	0.115	***
	COVID-19 incidence as of April 27th			0.713	***

Chi-square = 2.322; Degrees of freedom = 1; Chi-square/Degrees of freedom =2.322; p value = 0.128; Root mean square error of approximation = 0.045; Comparative fit index = 0.999.
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

July 8th.

5. Discussions

5.1. Discussions on influential factors

The results of our linear regression models and SEM not only verify certain influential factors discovered in the previous studies, but also disclose the significant association between COVID-19 incidence and several variables not yet well studied.

5.1.1. Reconfirmation of influential factors verified in existing literature

From the perspective of vulnerable population, nearly all the factors that have been repeatedly validated by existing literature present, once more, a statistically significant association with the COVID-19 infection in our models. Conforming with the existing literature, gender, age, ethnicity, income and smoking are significantly related to incidence. Gender differences in immunological function, smoking prevalence/pattern, employment behavior or willingness to engage in public health precautions (Chen et al., 2020; Liu et al., 2017; Park et al., 2010) may explain why the proportion of females is a variable significantly and negatively associated with COVID-19 incidence. Minor and older ages are associated with higher risk of infection, probably because of less rigorous immune responses, more comorbidities and underlying medical conditions in those age groups. Some African Americans might experience lower accessibility to affordable health insurance, and limited healthcare system capacity or investment (Millett et al., 2020), resulting in inferior health outcomes and high COVID-19 infection risk. Asians are found to be at lower risk of COVID-19 infection. Gould and Shierholz (2020) point out that this might be because more than a third of Asian Americans have the flexibility to work from home, while only one in five African Americans can adapt to a pattern of home-based work.

Income inequity (characterised by Income Ratio) presents significant association with higher incidence. This is consistent with existing literature highlighting the causal role of income inequity leading to negative health outcomes (Pickett & Wilkinson, 2015). Our analysis, however, finds that the 20th percentile of household income has a positive association with COVID-19 infection, which conflicts with the conclusion drawn by many existing studies (Borjas, 2020; Raifman & Raifman, 2020; Whittle & Diaz-Artiles, 2020). It does, however, conform with one earlier study, which identifies a strong positive correlation of COVID-19 incidence with median household income (Mollalo et al., 2020). One possible explanation is that limited testing resources at the early stage of the pandemic were disproportionately allocated to wealthier households. Another explanation might be that some high-income workers in certain occupations that require face-to-face communication are also at higher risk of COVID-19 infection. This suggests that policies during the pandemic should be developed not only to assist low-income population, but also to restrain some high-income groups facing potential high risks of virus exposure due to occupational demands.

With respect to transmission pathways, population density is positively associated with COVID-19 infection. Existing studies on the link between the pandemic and density have presented mixed results (Alirol et al., 2011; Borjas, 2020; Hamidi et al., 2020; Qi et al., 2020). Although our study provides new evidence for a positive effect of population density on COVID-19 transmission, it cannot be conclusive. The definition and measurement of density in different studies could influence the observed relationship between density and pandemic transmission. The Mobility Index defined in this paper presents a significantly positive association with COVID-19 incidence as of July 8th. This result is consistent with the existing literature that highlight the crucial effect of human mobility on disease transmission (Connolly et al., 2021; Shen, 2020).

5.1.2. Insights into potential influencing factors not yet well-explored

We examine a total of six potential influencing factors not yet well-explored and find that all of them associate significantly with the incidence of COVID-19.

The analytical results of unemployment rates show that counties with higher unemployment rates present a lower incidence of COVID-19. This could be because the unemployed commute less frequently, which, in turn, lowers exposure to pathogens.

This study also confirms that occupations categorized as Professional, Scientific and Technical Services bear a lower risk of COVID-19 infection. Transmission in these cases may be confined because these workers are able to work remotely and have lower interaction with the public during a public health emergency (Avdiu & Nayyar, 2020; Baker, 2020; Dingel & Neiman, 2020). However, the occupations categorized as Finance and Insurance are found to face higher risks, which suggests that these occupations may still require face-to-face interaction.

We believe the housing conditions might be relevant to transmission of communicable diseases and this conjecture has been verified. Overcrowding, higher housing costs, or lack of kitchen or plumbing facilities in dwellings impairs sanitary conditions and negatively affects infection prevention. Inferior plumbing facilities without disinfection and good sewage systems lead to the spread of the COVID-19 virus (WHO, 2020).

Counties with more neighborhood associations tend to have a higher incidence of COVID-19. This could be explained by the frequent gatherings and contacts in community social activities held by neighborhood

associations, increasing exposure to pathogens and possible secondary transmissions. Meanwhile, the positive correlation of Food Environment Index with COVID-19 infection indicates that better accessibility to healthy food may be associated with cluster transmission during shopping in crowded supermarkets.

In addition, counties with a higher ratio of population to Primary Care Physicians, meaning fewer Primary Care Physicians per capita, show significantly positive association with the number of confirmed COVID-19 cases. It is confirmed that adequate numbers of healthcare workers play an important role in providing timely detection, quarantine and treatment of COVID-19 patients to prevent further spread of the disease. This result is consistent with one presented in a recent paper stating that the proportion of nurse practitioners was negatively associated with COVID-19 infection rates (Mollalo et al., 2020).

5.2. Comparison between the two scenarios

Because the virus was not under control or conquered before the U.S. was reopened, COVID-19 cases resurged. Important issues can be identified in the discussions below by comparing the difference and linkage between the model results as of April 27th and July 8th for the two scenarios.

Following the reopening, people's activities and behavior changed dramatically. Dining inside restaurants and partying in bars resulted in spread of the virus. People engaged in more outdoor activities, especially when the air quality was good, which may have resulted in exposure to the virus and pandemic transmission. That may be why air quality plays a critical role in COVID-19 transmission in July. Meanwhile, it is not surprising to see that the Mobility Index became a positive and significant influential factor in July, because the majority of people were staying at home or nearby when the stay-at-home policy was imposed, but they socialised more following the reopening when face-to-face commerce was allowed to take place, leading to a higher risk of infection for those working in the Finance and Insurance industry. Our model captures these impacts. Certain premature reopening causes the unfortunate resurgence of confirmed cases.

COVID-19 incidence during the lockdown is found to have a strong causal linkage with the incidence after the opening. This can be explained by the fact that COVID-19 patients in the U.S. are not all quarantined and can still move around, leading to snowballing transmission over time. If the pandemic was already widespread during the lockdown in April, it can be more difficult to control it after the reopening in July, which emphasizes the importance of timely and thorough quarantining and treating of patients during the early stage of the pandemic.

Factors related to people's capacity to comply with the stay-at-home order directly affect COVID-19 incidence during the lockdown, but no longer have significant direct effects on incidence after the opening. Asian, unemployed people, and workers in Professional, Scientific, and Technical Services occupations were more capable of complying with the stay-at-home order during the lockdown and had a lower risk of infection. African-Americans and those on a low-income, however, may not be in a position to reduce their mobility during the lockdown due to engaging in occupations that prevent them from working from home (Baker, 2020), obliging them to face a higher infection risk. After the reopening, many people resumed their daily outdoor activities, making social distancing and other prevention policies, difficult, if not impossible, to follow. In these circumstances, people's ability to comply with the stay-at-home order is no longer a significant differentiating factor and the virus can attack people indiscriminately, resulting in everyone becoming vulnerable to the virus. In addition, with the reopening of small convenience stores and community grocery stores, some people would have multiple choices for purchase of healthy food rather than having to shop in crowded supermarkets, which explains why the Food Environment Index also does not have a significant direct effect on incidence as of July 8th.

6. Conclusion

The rapid spread of COVID-19 has become a major threat to the global public health systems and economies. To provide decision-making assistance in the fight against COVID-19, this paper explores COVID-19 incidence and three aspects of its associated influential factors based upon datasets collected from high-incidence counties in the United State during the first wave of the pandemic. A research framework and experimental results are presented. In addition to factors mentioned in the existing literature, we uncover several significant variables not yet widely recognized, including unemployment rate, the percentage of households with severe housing problems, neighborhood association rate, Food Environment Index, the percentage of employment in Professional, Scientific and Technical services as a home-based occupation, as well as ratio of population to Primary Care Physicians. Based on the datasets from different scenarios, the results provide insight into ways to make effective decisions to subdue the virus and mitigate the infectious risk.

For all counties and regions throughout the world, we recommend that attention should focus on changes in groups of vulnerable populations during the lockdown and after the reopening, so that timely support can be made available. During the lockdown, there is a particular need to protect and assist vulnerable people who are less capable of complying with the stay-at-home order because of their occupations or socio-economic disadvantage, including minorities, those on a low-income and others in occupations less amenable to home-based work. After the reopening, workers in the Finance and Insurance industry should receive attentions. Those in the under 18 and over 65 age groups need to be protected both during the lockdown and after the reopening. To protect vulnerable populations and ensure socio-economic stability during such crises, social protection schemes (e.g., medical insurance, relief grants, vaccination supplies) should be in place to help people cope with economic uncertainties or virus exposure. Policies should be developed to provide additional workplace protection and to initiate more jobs that can be conducted from home with flexible hours.

To cut off the transmission chain, it is important to quarantine and treat patients timely, maintain social distancing after the reopening, and issue additional policies focussing on housing conditions, healthy food and neighborhood associations. Timely and thorough quarantining and treating of patients is essential to avoid the snowballing transmission of COVID-19 over time. Following a long lockdown, people are eager to go out, especially when the air quality is good, and it is necessary to remind them to take personal protection and maintain social distancing while outdoors. In addition, special policies and planning should be made to improve neighborhoods with unsanitary and overcrowded housing to reduce the risk of cluster transmission in high-density environments. This is particularly important in countries and regions with high population density and large numbers of overcrowded housing. Policies are also needed to ensure that people have safe access to healthy food during a pandemic, preferably by providing a contactless home delivery service of healthy food, or by strictly controlling customer flow in the supermarkets. Neighborhood associations should be warned of the importance of minimising face-to-face community social activities during an outbreak.

Finally, the healthcare workforce plays a crucial role in lowering contagious risk and preventing the wide spread of the disease. Governments with limited healthcare resources should improve and invest more on their public health systems to ensure that adequate healthcare resources are available in the event of a pandemic outbreak.

This study on the influential factors of county-level COVID-19 incidence in the United States presents early findings and proves the effects and significance of lockdown. The above policy recommendations, based on these findings, are critical for stopping an outbreak in the early stages before the pandemic gets out of control, and provide suggestions for mitigating the negative effects of subsequent waves of the pandemic. It also brings clues to develop diverse supportive policies to protect the

public from the virus after the reopening. As the COVID-19 pandemic continues to spread around the globe and the number of deaths and infections increase daily, the factors identified in this study and their corresponding policies may provide referable solutions for other countries in their ongoing fight against the pandemic.

CRediT authorship contribution statement

Lan Wang: Conceptualization, Methodology, Validation, Writing - review & editing. **Surong Zhang**: Literature review, Data collection, Writing - original draft. **Zilin Yang**: Literature review, Writing - original draft. **Ziyu Zhao**: Data collection, Formal analysis, Writing - original draft. **Anne Vernez Moudon**: Conceptualization, Writing - review & editing. **Huasen Feng**: Data collection, Formal analysis. **Junhao Liang**: Data collection, Formal analysis. **Wenyao Sun**: Writing - review & editing. **Buyang Cao**: Conceptualization, Methodology, Validation, Writing - review & editing.

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.

Acknowledgments

This research is supported by National Natural Science Foundation of China (41871359, 52078349).

References

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases, 11*(2), 131–141.

American Journal of Managed Care. (2020). A timeline of COVID-19 developments in 2020. ajmc.com/focus-of-the-week/a-timeline-of-covid19-developments-in-2020.

Avdiu, B., & Nayyar, G. (2020). When face-to-face interactions become an occupational hazard: Jobs in the time of COVID-19. In *, 9240. World Bank Policy Research Working Paper*.

Babey, S. H., Diamant, A. L., Hastert, T. A., & Harvey, S. (2008). *Designed for disease: The link between local food environments and obesity and diabetes*. UCLA Center for Health Policy Research.

Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modeling study. *The Lancet Infectious Diseases, 20*(11), 1247–1254.

Baker, M. G. (2020). Characterizing occupations that cannot work from home: a means to identify susceptible worker groups during the COVID-19 pandemic. *MedRxiv*.

Baker, M. G., McDonald, A., Zhang, J., & Howden-Chapman, P. (2013). Infectious diseases attributable to household crowding in New Zealand: a systematic review and burden of disease estimate. In *He Kainga Oranga/ Housing and Health Research Programme*. Wellington: University of Otago.

Becker, S., & Soukup, J. M. (1999). Exposure to urban air particulates alters the macrophage-mediated inflammatory response to respiratory viral infection. *Journal of Toxicology and Environmental Health Part A, 57*(7), 445–457.

Behavioral Risk Factor Surveillance System. (2017). Health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/health-behaviors/tobacco-use/adult-smoking.

Borjas, G. J. (2020). *Demographic determinants of testing incidence and COVID-19 infection in New York City neighborhoods (No. w26952)*. National Bureau of Economic Research.

Bureau of Labor Statistics. (2018). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/social-and-economic-factors/employment/unemployment.

Cai, H. (2020). Sex difference and smoking predisposition in patients with COVID-19. *The Lancet Respiratory Medicine, 8*(4), Article e20.

Census Population Estimates. (2016). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/2020-measures.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., & Yu, T. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet, 395*(10223), 507–513.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). *Labor markets during the covid-19 crisis: A preliminary view (No. w27017)*. National Bureau of Economic Research.

Comprehensive Housing Affordability Strategy (CHAS). (2016). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/2020-measures.

Connolly, C., Keil, R., & Ali, S. H. (2021). Extended urbanisation and the spatialities of infectious disease: demographic change, infrastructure and governance. *Urban Studies, 58*(2), 245–263, 0042098020910873.

County Business Patterns. (2017). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/social-and-economic-factors/family-social-support/social-associations.

Dingel, J. I., & Neiman, B. (2020). *How many jobs can be done at home? (No. w26948)*. National Bureau of Economic Research.

Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., & Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences, 117*(18), 9696–9698.

Environmental Public Health Tracking Network. (2014). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/physical-environment/air-water-quality/air-pollution-particulate-matter.

ESRI. (2020). COVID-19 overview. https://coronavirus-resources.esri.com/datasets/1044bb19da8d4dbfb6a96eb1b4ebf629_0/data?geometry=80.683%2C-16.820%2C-105.645%2C72.123.

Fang, L., Karakiulakis, G., & Roth, M. (2020). Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *The Lancet Respiratory Medicine, 8*(4), e21.

ii Glanz, K., Sallis, J. F., Saelens, B. E., & Frank, L. D. (2005). Healthy nutrition environments: Concepts and measures. *American Journal of Health Promotion: AJHP, 19*, 330–333.

Gould, E., & Shierholz, H. (2020). Not everybody can work from home: Black and Hispanic workers are much less likely to be able to telework. Economic Policy Institute. https://www.epi.org/blog/black-and-hispanic-workers-are-much-less-likely-to-be-able-towork-from-home/. (Accessed 16 April 2021).

Grasselli, G., Pesenti, A., & Cecconi, M. (2020). Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: Early experience and forecast during an emergency response. *JAMA, 323*(16), 1545–1546.

Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., & Zhao, L. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes/Metabolism Research and Reviews, e3319*.

Hamidi, S., Sabouri, S., & Ewing, R. (2020). Does density aggravate the COVID-19 pandemic? Early findings and lessons for planners. *Journal of the American Planning Association, 86*(4), 495–509.

Ioannidis, J. P., Axfors, C., & Contopoulos-Ioannidis, D. G. (2021). Second versus first wave of COVID-19 deaths: Shifts in age distribution and in nursing home fatalities. *Environmental Research, 195*, Article 110856.

Johns Hopkins University GitHub open-source project. (2020). https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series.

Kass, D. A., Duggal, P., & Cingolani, O. (2020). Obesity could shift severe COVID-19 disease to younger ages. *The Lancet (London, England), 395*(10236), 1544–1545.

Kawohl, W., & Nordt, C. (2020). COVID-19, unemployment, and suicide. *The Lancet Psychiatry, 7*(5), 389–390.

Kraemer, M. U., Yang, C. H., Gutierrez, B., Wu, C. H., Klein, B., Pigott, D. M., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science, 368*(6490), 493–497.

Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Ichii, H., Zacharksi, M., & Khosrawipour, T. (2020). The association between international and domestic air traffic and the coronavirus (COVID-19) outbreak. *Journal of Microbiology, Immunology and Infection, 53*(3), 467–472.

Leddy, A. M., Weiser, S. D., Palar, K., & Seligman, H. (2020). A conceptual model for understanding the rapid COVID-19–related increase in food insecurity and its impact on health and healthcare. *The American Journal of Clinical Nutrition, 112*(5), 1162–1169.

Li, R., Rivers, C., Tan, Q., Murray, M. B., Toner, E., & Lipsitch, M. (2020). The demand for inpatient and ICU beds for COVID-19 in the US: Lessons from Chinese cities. *MedRxiv*.

Lighter, J., Phillips, M., Hochman, S., Sterling, S., Johnson, D., Francois, F., & Stachel, A. (2020). Obesity in patients younger than 60 years is a risk factor for Covid-19 hospital admission. *Clinical Infectious Diseases, 9*(10.1093).

Liu, S., Zhang, M., Yang, L., Li, Y., Wang, L., Huang, Z., & Zhou, M. (2017). Prevalence and patterns of tobacco smoking among chinese adult men and women: findings of the 2010 national smoking survey. *Journal of Epidemiology and Community Health, 71*(2), 154–161.

Luo, W., Majumder, M., Liu, D., Poirier, C., Mandl, K., Lipsitch, M., & Santillana, M. (2020). *The role of absolute humidity on transmission rates of the COVID-19 outbreak*.

Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., & Sherwood, J. (2020). Assessing differential impacts of COVID-19 on black communities. *Annals of Epidemiology, 47*, 37–44.

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, 138884*.

Morland, K. B., & Evenson, K. R. (2009). Obesity prevalence and the local food environment. *Health & Place, 15*(2), 491–495.

Novel, C. P. E. R. E. (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi, 41*(2), 145.

Pan, A., Liu, L., Wang, C., Guo, H., Hao, X., Wang, Q., & Wei, S. (2020). Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA, 323*(19), 1915–1923.

Pareek, M., Bangash, M. N., Pareek, N., Pan, D., Sze, S., Minhas, J. S., & Khunti, K. (2020). Ethnicity and COVID-19: an urgent public health research priority. *The Lancet, 395*(10234), 1421–1422.

Park, J. H., Cheong, H. K., Son, D. Y., Kim, S. U., & Ha, C. M. (2010). Perceptions and behaviors related to hand hygiene for the prevention of H1N1 influenza transmission among korean university students during the peak pandemic period. *BMC Infectious Diseases, 10*(1), 222.

Pickett, K. E., & Wilkinson, R. G. (2015). Income inequality and health: a causal review. *Social Science & Medicine, 128*, 316–326.

Qi, C. F., Yang, L. R., Yang, Z. X., Shang, L., Xie, G. L., Wang, R. Q., Wang, S. S., & Yang, W. F. (2020). The influencing factors of interprovincial transmission and development of COVID-19: Data analysis based on 30 provinces and cities in China [2020-07-22] *Journal of Xi'an Jiaotong University (Medical Sciences)*, 1–13 http://kns.cnki.net/kcms/detail/61.1399.r.20200417.1413.002.html.

Raifman, M., & Raifman, J. (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.

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

Shen, J. (2020). Covid-19 and inter-provincial migration in China. *Eurasian Geography and Economics, 61*(4–5), 620–626.

Sinclair, A. J., & Abdelhafiz, A. H. (2020). Age, frailty and diabetes–triple jeopardy for vulnerability to COVID-19 infection. *EClinicalMedicine*, 22.

Smith, K. F., Dobson, A. P., McKenzie, F. E., Real, L. A., Smith, D. L., & Wilson, M. L. (2005). Ecological theory to enhance infectious disease control and public health policy. *Frontiers in Ecology and the Environment, 3*(1), 29–37.

Starfield, B., Lemke, K. W., Bernhardt, T., Foldes, S. S., Forrest, C. B., & Weiner, J. P. (2003). Comorbidity: implications for the importance of primary care in 'case'management. *The Annals of Family Medicine, 1*(1), 8–14.

Stier, A., Berman, M., & Bettencourt, L. (2020). COVID-19 attack rate increases with city size. In *Mansueto Institute for Urban Innovation Research Paper Forthcoming*.

Travaglio, M., Yu, Y., Popovic, R., Selley, L., Leal, N. S., & Martins, L. M. (2021). Links between air pollution and COVID-19 in England. *Environmental Pollution, 268*, Article 115859.

U.S. Census American Community Survey. (2018). Open data network. https://www.opendatanetwork.com.

U.S. Department of Commerce. (2020). Bureau of economic analysis (BEA). https://apps.bea.gov/regional/downloadzip.cfm.

U.S. Department of Transportation. (2020). https://data.bts.gov/Research-and-Statistics/Trips-by-Distance/w96p-f2qv.

United States Diabetes Surveillance System. (2016). <article-title>County health ranking & roadmaps</article-title>. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/health-behaviors/diet-exercise/adult-obesity.

USDA Food Environment Atlas. (2015). County health ranking & roadmaps. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/county-health-rankings-model/health-factors/health-behaviors/diet-exercise/limited-access-to-healthy-foods.

Vardavas, C. I., & Nikitara, K. (2020). COVID-19 and smoking: A systematic review of the evidence. *Tobacco Induced Diseases, 18*.

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., & Zhao, Y. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–Infected pneumonia in Wuhan, China. *JAMA, 323*(11), 1061–1069.

Whittle, R. S., & Diaz-Artiles, A. (2020). An ecological study of socioeconomic predictors in detection of COVID-19 cases across neighborhoods in New York City. *BMC Medicine, 18*(1), 1–17.

World Health Organization. (2018). WHO housing and health guidelines. https://www.who.int/publications/i/item/who-housing-and-health-guidelines. (Accessed 22 July 2020).

World Health Organization. (2020). Strengthening preparedness for COVID-19 in cites and urban settings: interim guidance for local authorities. https://www.who.int/publications/i/item/strengthening-preparedness-for-covid-19-in-cities-and-urban-settings.

Wu, C., Chen, X., Cai, Y., Zhou, X., Xu, S., Huang, H., & Song, J. (2020). Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Internal Medicine, 180*(7), 934–943.

Wu, Z., & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *Journal of the American Medical Association, 323*(13), 1239–1242.

Xu, H., Yan, C., Fu, Q., Xiao, K., Yu, Y., Han, D., & Cheng, J. (2020). Possible environmental effects on the spread of COVID-19 in China. *Science of the Total Environment, 731*, Article 139211.

Xu, Q., Li, X., Wang, S., Wang, C., Huang, F., Gao, Q., & Guo, X. (2016). Fine particulate air pollution and hospital emergency room visits for respiratory disease in urban areas in Beijing, China, in 2013. *PLoS One, 11*(4), Article e0153099.

Yang, Y., Li, W., Zhang, Q., Zhang, L., Cheung, T., & Xiang, Y. T. (2020). Mental health services for older adults in China during the COVID-19 outbreak. *The Lancet Psychiatry, 7*(4), Article e19.

YouTube. The Economist. (2020). Covid-19: what you need to know about the second wave. https://www.youtube.com/watch?v=uOY32cU0ePc.

Zhang, Z., Xue, T., & Jin, X. (2020). Effects of meteorological conditions and air pollution on COVID-19 transmission: evidence from 219 chinese cities. *Science of the Total Environment, 741*, Article 140244.

Zheng, Z., Peng, F., Xu, B., Zhao, J., Liu, H., Peng, J., & Ye, C. (2020). Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis. *Journal of Infection, 81*(2), e16–e25.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., & Guan, L. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet, 395*(10229), 1054–1062.

Zhu, Y., Xie, J., Huang, F., & Cao, L. (2020). Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Science of the Total Environment, 727*, Article 138704.