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.

Environmental Research 202 (2021) 111657

Contents lists available at ScienceDirect

Environmental Research

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

COVID-19 in Asia: Transmission factors, re-opening policies, and vaccination simulation

Maryam Baniasad [a], Morvarid Golrokh Mofrad [b], Bahare Bahmanabadi [c], Sajad Jamshidi [d,*]

[a] Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, 43210, USA
[b] Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran
[c] Department of Water Engineering, Imam Khomeini International University, Qazvin, Iran
[d] Department of Agronomy, Purdue University, West Lafayette, IN, 47907, USA

ARTICLE INFO

Keywords:
Air pollution
COVID-19
Contact tracing
Deep learning
Government stringency index
Meteorological variables
PM2.5
Pandemic
Vaccination

ABSTRACT

This work aims to provide insights on the COVID-19 pandemic in three prime aspects. First, we attempted to understand the association between the COVID-19 transmission rate, environmental factors (air pollution, weather, mobility), and socio-political parameters (Government Stringency Index, GSI). Second, we evaluated the efficiency of various strategies, including radical opening, intermittent lockdown, phase lift, and contact tracing, to exit the COVID-19 pandemic and get back to pre-pandemic conditions using a stochastic individual-based epidemiology model. Third, we used a deep learning approach and simulated the vaccination rate and the time for reaching herd immunity. The analysis was done based on the collected data from eight countries in Asia, including Iran, Turkey, India, Saudi Arabia, United Arab Emirates, the Philippines, South Korea, and Russia (as a transcontinental country). Our findings in the first part highlighted a noninfluential impact from the weather-driven parameters and short-term exposure to pollutants on the transmission rate; however, long-term exposure could potentially increase the risk of COVID-19 mortality rates (based on 1998–2017 p.m.2.5 data). Mobility was highly correlated with the COVID-19 transmission and based on our causal analysis reducing mobility could curb the COVID-19 transmission rate with a 6-day lag time (on average). Secondly, among all the tested policies for exiting the COVID-19 pandemic, the contact tracing was the most efficient if executed correctly. With a 2-day delay in tracing the virus hosts, a 60% successful host tracing, and a 70% contact reduction with the hosts, a pandemic will end in a year without overburdening a healthcare system with 6000 hospital beds capacity per million. Lastly, our vaccine simulations showed that the target date for achieving herd immunity significantly varied among the countries and could be delayed to October–november 2022 in countries like India and Iran (based on 60% immunized population and assuming no intermediate factors affecting the vaccination rate).

1. Introduction

The novel human coronavirus (2019-nCoV), also known as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), first appeared in late 2019 in Wuhan Province, China (Zhu et al., 2020). As of March 23rd, 2021, more than 123 million cases and 2.7 million deaths have been reported due to the COVID-19 worldwide (WHO COVID-19 Dashboard. Geneva: World Health Organization, 2020. Available online: https://covid19.who.int/). Coronaviruses belong to the subfamily of *Orthocoronavirinae* in the *Coronaviridae* family. The viral particles consist of a positive-sense, single-stranded RNA (+ssRNA), an enveloped capsid, and six functional open reading frames. Currently, this family has four genera α, β, γ, and δ, and SARS-CoV-2, like the SARS-CoV-1 and MERS-CoV, belongs to the genus β (Zhou et al., 2020). Many clinical and non-clinical (population-size) studies have considered evaluating different aspects of the COVID-19 pandemic (Cheng et al., 2020b; Lotfi et al., 2020; Wang et al., 2020). Re-evaluating the analysis at different scales with more data becomes available through the pandemic helps

Abbreviations: Open reading frame, (ORF); Severe acute respiratory syndrome, (SARS); SARS coronavirus 2, (SARS-CoV-2); Middle east respiratory syndrome, (MERS); Novel coronavirus 2019, (2019-nCoV); Coronavirus disease-2019, (COVID-19); Global positioning system, (GPS); Ministry of health and medical education, (MHME); GISAID, (Global initiative on sharing all influenza data); Particulate matter less than 2.5 μm in diameter, (PM2.5); Sulfur dioxide, (So2); Temperature, (Temp); Government stringency index, (GSI); Relative humidity, (RH); Respiratory syncytial virus, (RSV).

* Corresponding author. Research assistant at Purdue University, Department of Agronomy, West Lafayette, IN, 47907, USA.
E-mail address: Sjamshi@purdue.edu (S. Jamshidi).

https://doi.org/10.1016/j.envres.2021.111657
Received 25 March 2021; Received in revised form 22 June 2021; Accepted 30 June 2021
Available online 8 July 2021
0013-9351/

validate/reject the early-stage studies' findings. Additionally, with the manufacturing and development of the COVID-19 vaccine, we need to consider future scenarios to exit the pandemic and go back to normal conditions effectively.

COVID-19 is a respiratory infectious disease and can survive in the aerosol for hours; thus, it is essential to investigate the association between air pollution and COVID-19 infection to identify potential exposure risk (Zhang et al., 2020). Recent studies on the COVID-19 in different countries have examined the relationship between air pollution and the incidence and mortality of the virus. Based on some of the studies, geographical patterns of COVID-19 transmission and mortality among countries are linked with local levels of pollutants (Conticini et al., 2020; Frontera et al., 2020; Singh et al., 2021). Areas with limited health facilities, high population density, air pollution, and vulnerable climatic conditions are at higher risk of the virus spread. For example, Salam (2020) studied air quality indicators (PM2.5, PM10, AQI, and nitrogen dioxide (NO_2)), carbon dioxide (CO_2) emissions, and climate variables (temperature, relative humidity, rainfall, and wind) in Dhaka, Bangladesh, and found a significant positive correlation between COVID-19 cases and air quality indices and climatic variables. Nevertheless, Wu et al. (2020) reported that the association between air pollution needs to be re-evaluated using long-term exposure data.

Meteorological variables have also been investigated for their impact on the COVID-19 transmission and severity. Some studies have reported that factors such as temperature and humidity could impact the COVID-19 transmission rate (Damette and Goutte, 2020; Rahman et al., 2020; Rizwanul Fattah et al., 2020; Sajadi et al., 2020). For example, Ma et al. (2020) reported that diurnal temperature and humidity change affect COVID-19 infections in Wuhan, China. On the other hand, studies like Jamshidi et al. (2020) reported no compelling evidence for including weather as a significant contributor to the spread of the COVID-19. According to this study, a higher level of mobility and traveling can increase the spread of the disease and transfer it from regional levels to global epidemics. The inconsistencies between the study's findings highlight the importance of re-considering these factors for a larger time window and over different spatial scales.

Many studies confirm the importance of screening, surveillance, and control efforts, particularly at airports and other transportation hubs, to prevent the spread of SARS-CoV-2 (Imai et al., 2020; Kraemer et al., 2020; Stoecklin et al., 2020; Tian et al., 2020). Accordingly, authorities have imposed various lockdown or tracing policies to curb the COVID-19 pandemic growth. While executing these policies has been associated with lowering the risk of the transmission rate, the impact of lifting the limitations and re-opening remains unclear (Cheng et al., 2020a). de Vlas and Coffeng (2021) evaluated the impact of different policies on COVID-19 transmission for exiting the COVID-19 pandemic in the Netherland. They proposed a "phased lift of control" strategy as a manageable and effective strategy to reach herd immunity (without considering the impact of vaccination). Kraemer et al. (2020) performed a study using real-time mobility from Wuhan and detailed case data, including travel history, to show the role of transmission case importation in cities across China and showed the impact of control measures. This study shows that China's drastic control measures substantially mitigated the spread of COVID-19.

Vaccine development is another aspect that can significantly affect the policies countries are imposing against COVID-19 for ending the pandemic. As of the time of developing this study (March 2021), most countries are in early-stage vaccinations. Several studies have reported the significant impact of the vaccination on reducing the COVID-19 transmission and prioritizing vaccine allocations (Bubar et al., 2021; Mukandavire et al., 2020; Yang et al., 2020). The information on vaccine development and predicting the future scenarios based on the current vaccination rollout speeds could significantly contribute to incorporating more effective measures to confront COVID-19 in different countries.

The current study aims to identify the impact of environmental and socio-political factors on COVID-19 transmission in eight countries, including Iran, Turkey, India, Russia, Saudi Arabia, the United Arab Emirates, the Philippines, and South Korea. Russia is a transcontinental country extending from Asia to Europe, with its most land in Asia and most population in Europe. Yet, to be consistent with the context, we refer to it as in Asia. Next, we evaluated the impact of different lockdown/re-opening strategies on the transmission rates to find the optimal approach for returning to normal (pre-pandemic) conditions. Comparing these factors and the policies can guide the public and governments in controlling COVID-19 transmission more effectively during similar critical situations. Therefore, in this study, we have also compared the vaccination speed in these countries during the initial vaccination phase. The negative binomial regression was used to forecast vaccination progress and the amount of time needed to reach vaccine-induced herd immunity, considering the population and current vaccination speed for each country.

2. Materials and methods

2.1. Study areas

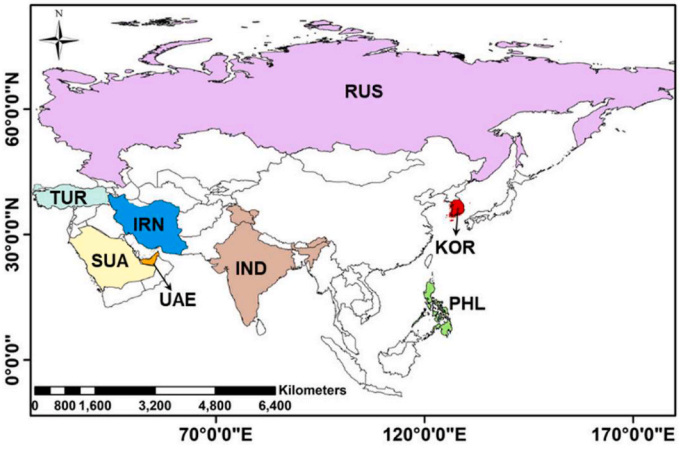
We analyzed eight Asian countries, including Iran, Turkey, Saudi Arabia, United Arab Emirates, India, Russia, the Philippines, and South Korea. The selection of these countries was made based on two prime criteria. First, these countries were selected to represent various geospatial and political characteristics (e.g., different climates, socio-economic conditions). Second, we considered the availability of the data required for the analysis (discussed in the following subsection). Fig. 1 shows the locations of the selected countries on the map with their demographic information. The population of the selected countries ranges from 9.9 million (United Arab Emirates) to 1.380 billion (India), with the variable climate ranging from arid (Saudi Arabia) to tropical (Philippines), temperate (North of India), and very cold (Russia).

The analysis was done in three main sections. In the first part, we performed a correlation and causality analysis between the COVID-19 infection rate and a number of environmental and social factors, including weather (air temperature and humidity), air pollution, mobility, and governmental responses. The objective of this analysis was to understand the driver factors of COVID-19 at a large scale. In the second part of the analysis, we evaluated the impact of different strategies for getting back to pre-pandemic (normal) conditions using a stochastic individual-based epidemiology model (the VirSim model). The objective of this section was to shed light on the success of different exit strategies that countries could adopt. Lastly, for the third part, we compared the vaccination progress of these eight countries using the daily administered doses of vaccine during the initial vaccination phase (Dec 2020–Mar 2021). The objective of this section is to highlight the impact of vaccination speed on the ultimate control of COVID-19 by predicting the amount of time needed to vaccinate 60% of each country's population considering the current daily average of administered doses.

2.2. Data collection

We collected daily data of COVID-19 infection and death rate from March to December 2020 from the World Health Organization (available at https://www.who.int/). A subset of the data associated with the eight countries, selected for the evaluations, was furthered processed (i.e., quality checked, converted to desired formats for modeling purposes).

Daily meteorological variables, including the air temperature (Tair) and relative humidity (Rh), were retrieved from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Gelaro et al., 2017). We used MERRA-2 data because it combines disparate meteorological measurements and generates physically uniform and spatially consistent gridded datasets at hourly time



Country	Population	Land Area (km^2)	Urban (%)	Density (people/km^2)
India	1,380,004,385	2,973,190	35%	464
Philippines	109,581,078	298,170	47%	368
Turkey	84,339,067	769,630	76%	110
Iran	83,992,949	1,628,550	76%	52
South Korea	51,269,185	97,230	82%	527
Saudi Arabia	34,813,871	2,149,690	84%	16
United Arab Emirates	9,890,402	83,600	86%	118
Russia	145,934,462	16,376,870	74%	9

Fig. 1. Map of the study area (Asia) with locations of the countries evaluated in this study, including Iran (IRN), Turkey (TUR), Russia (RUS), India (IND), Saudi Arabia (SAU), United Arab Emirates (UAE), Philippines (PHL), and South Korea (KOR) and their demographic information.

steps and spatial resolution of ~50 km. The Tair and Rh data were spatially averaged for each of the countries, temporally reaggregated to daily, and used in the analysis.

To evaluate the interplay between air quality and COVID-19 infection risk, we considered a short and long-term analysis. First, we performed a short-term analysis by evaluating the association between the satellite-based AOD (Aerosol Optical Depth) and COVID-19 cases during the pandemic (March to December 2020). AOD is a measure of the extinction of the solar beam by particles in the atmosphere, and thus, provides a decent metric of air quality. We should highlight that air quality should be ideally represented using PM2.5 data (fine particle matter with a diameter smaller than 2.5 μm), however, the satellite-driven PM2.5 data were not available during the COVID-19 pandemic, and thus we used the AOD data instead. Many studies have reported a strong correlation between the AOD and PM2.5 (Jin et al., 2020; Krishna et al., 2019). The AOD data were retrieved from the Sentinel 3 (A and B) satellites available from Copernicus Climate Data Store (https://cds.climate.copernicus.eu/). The data has been retrieved at a daily level and spatially averaged over each of the countries for further analysis.

Second, we considered a long-term analysis to evaluate the association between the long-term exposure to pollutants (using PM2.5 data) and susceptibility to COVID-19 infection. For the long-term analysis, we used all available PM2.5 data from 1998 to 2017 from the Socioeconomic Data and Applications Center (SEDAC). The PM2.5 data were temporally averaged for all the years, and the long-term averaged data were used to evaluate the link between the hotspots of air pollution and COVID-19 mortality rates.

The mobility data were retrieved from the GitHub repository of the reports generated by Google, Apple, and Waze COVID-19 Community Mobility Reports (https://github.com/ActiveConclusion/COVID19_mobility). A baseline was calculated for each weekday by averaging the values during February (as the pre-pandemic phase). The daily percentage change from the baseline corresponding to the weekdays was calculated and used as the mobility index (MI). The MI demonstrates the percentage of increase (positive values) or decreases (negative values) in the daily trips from March to December 2020 with reference to the baseline (during February).

The government stringency index (GSI) is a composite measure based on nineteen response indicators, including eight indicators on containment and closure policies, four economic policies, and seven indicators of health system policies. The index is calculated as the metrics mean score, each taking a value between 0 and 100 (i.e., 100 = strictest response). These data were provided by the Oxford COVID-19 Government Response Tracker (https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker#data), and we used a subset of the data for the countries used in the analysis.

2.3. Models and simulations

For simulating the strategies to exit the COVID-19 pandemic, we used the VirSim model (de Vlas and Coffeng, 2021) and considered four different scenarios. VirSim is a compartmental SEIR (Susceptible, Exposed, Infected, Recovered) model that considers a region or country by clustering individuals into small groups (e.g., sub-city scale, villages) with each group being a part of a supercluster (e.g., city or province). The model details and its technical aspects can be found at the "science versus corona" project (available from https://scienceversuscorona.com/interactive-exploration-of-covid-19-exit-strategies/), but in essence, the model explores the efficiency of different policies for controlling the COVID-19 pandemic growth and exiting the pandemic. The simplified schematic of the model workflow is presented in Fig. 2. The model calculates the force of infection (λ) as a function of i) the virus infection's occurrence in the cluster and superclusters, ii) variations of the transmission rate and contact rate, and iii) the effect of interventions considered by the model.

The VirSim model used in the study simulates an epidemic by randomly seeding 10 infections. The model then considers a general lockdown condition for 60 days (7 days of loose lockdown and 53 days of stricter lockdown) when the number of infections extends a certain level. The interventions are considered by the model 30 days later. These values are the default values considered by the developers of the models, yet, since we intended to compare the different strategies, we kept these values unchanged and similar among the different simulations. The model parameterizations are provided in Supplementary Table S1.

With the VirSim model, we have tested the impact of four strategies

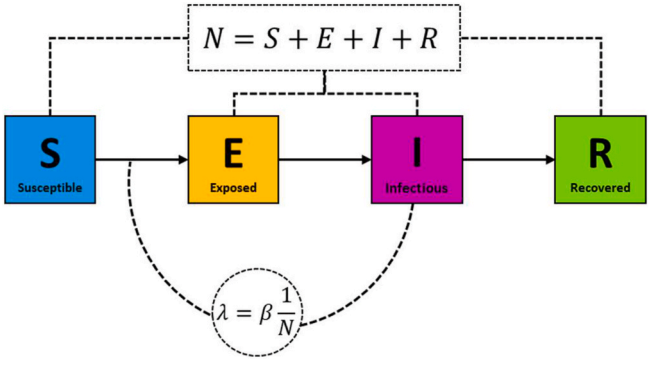


Fig. 2. The simplified graphical representation of SEIR (susceptible, Exposed, Infectious, Recovered) model. λ is the force of infection, β is the average contact rate, and N is the population size.

for ending the semi-lockdown conditions that countries have been regulating. The first tested policy is the Radical Opening (RO), in which all the businesses will get back to a normal condition (pre-pandemic) with a sudden lift of all imposed policies. The second tested policy is the Phased Lift of Control (PLoC), in which a country lifts the lockdown in different areas at different times. The third tested policy is the intermittent lockdown, in which several cycles of lockdown occur (a lockdown is imposed whenever the COVID-19 cases increase to a particular level, and the lockdown is removed after a certain time (considered as 25 days in the simulations). The last tested scenario is contact tracing, in which the infectious/symptomatic cases are identified and isolated to reduce the risk of transferring the virus. The variations of these scenarios are evaluated for a country with ten regions, including an equal number of sub-regions and 1000 individuals per sub-regions.

2.4. Analysis and statistics

For evaluating the associations between the individual variables and the COVID-19 cases, we first applied generalized linear models (GLM) to account for the non-linear behavior of the variables and their non-normal distributions. In the GLM models, each outcome (Y) of the dependent variables (x) for each observation (i) is assumed to be:

$$Y_i \sim F_{EDM}(\cdot|\theta,\varphi,\omega_i)\&\mu_i = E[Y_i|x_i] = g^{-1}\left(x_i'\beta\right) \quad (1)$$

where F_{EDM} denotes a distribution family, μis the mean of the distribution, g is the link function, $x_i'\beta$ is the linear predictor, and $E[Y_i|x_i]$ is the expected value of Y conditional on x. Based on the resulting distribution of individual variables, we applied various GLM including the Poisson, Binomial, and Gaussian models. The robustness, significance, and relative quality of the applied models and correlations were examined by the chi-square $(\varkappa^2)$, P_{value}, and Akaike information criterion (AIC), respectively. A superior model performance is linked with a lower magnitude of $\varkappa^2$ and AIC.

Given the time-dependent and non-linear nature of the real-world data, as denoted by Jamshidi et al. (2020), the correlation-based analysis could not reveal and quantify the causal relationships between the observations. For example, a number of studies have reported on a strong correlation between COVID-19 mortality rate and air pollution, but the question remains if the COVID-19 and its restrictions caused the change in the air pollution or changes in air pollution have caused the COVID-19 cases to rise or drop. Therefore, the interconnection between the variables was evaluated through a causality analysis.

We identified the time-lagged causal dependency structure (considering ten time-lagged steps) underlying the networks of the variables (Tair, Rh, MI, GSI, and AOD) and quantified causal strength using the PCMCI index (Runge et al., 2019). The PCMCI index considers a time-dependent network of variables $(X_t^1, X_t^2, \ldots, X_t^N)$ with:

$$X_t^j = f_j\left(P\left(X_t^j\right), \eta_t^j\right) \quad (2)$$

where f_j is the non-linear function between the variables, η_{jt} is the noise, and $P(X_{jt})$ denotes the causal parents of variable X_{jt} at different time lags $(X_{t-1}, X_{t-2}, \ldots)$. Next, it applies the principal component (PC) using a Markov set discovery through an iterative process to only select the relevant parents and conditions, i.e., $\widehat{P}(X_{jt})$, for all the variables as:

$$\mathrm{PC}: X_{t-\tau}^j \not\perp\!\!\!\perp X_t^j | \mathrm{P} \text{ for any P with } |\mathrm{P}| = \mathrm{p} \quad (3)$$

Then, the model employs the momentary conditional independence (MCI) test to capture the dependencies at lag-time steps:

$$\mathrm{MCI} = X_{t-\tau}^i \not\perp\!\!\!\perp X_t^j | \widehat{\mathrm{P}}\left(X_t^j\right) \backslash \{, \widehat{\mathrm{P}}\left(X_{t-\tau}^i\right) \quad (4)$$

We performed the causality tests and calculated the PCMCI index using the Tigramite open-source software package (https://github.com/jakobrunge/tigramite). The dependency structure of the variables is represented in a graphical model to make it visually simpler to understand.

In the last part of the manuscript, we used a deep learning approach combined with the negative binomial regression model (NBR) (Cameron and Trivedi, 2013) to predict the number of administered COVID-19 vaccine doses and an approximation date that countries reach vaccine-based herd immunity. For the deep learning approach, we used the Long Short Term Memory network (LSTM) approach (Hochreiter and Schmidhuber, 1997). We selected this model primarily due to its high capability of learning long-term dependencies and capturing the patterns in sequences of data, such as numerical time's series data.

LSTM takes the current and historical inputs and applies three sigmoid layers (known as "gates") to decide which information to be used or removed from the model. The gates include a forget gate (decides what information to be removed), an input gate (decides which information to update), and, and an output gate (decides which information to be generated as the output). The mathematical form of these gates is summarized below.

$$f_t = \sigma\left(W_f \times [h_{t-1}, x_t] + b_f\right) \quad (5)$$

$$i_t = \sigma(W_i \times [h_{t-1}, x_t] + b_i) \quad (6)$$

$$\widetilde{C}_t = tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (7)$$

$$C_t = C_{t-1} \times C_t + i_t \times \widetilde{C}_t \quad (8)$$

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_0) \quad (9)$$

$$h_t = o_t \times tanh(C_t) \quad (10)$$

where f_t is the forget gate vector, i_t is the input/update gate vector, h_t is the output gate vector, σ is the sigmoid function, W is the weight matrices, and c_t is the cell state vector.

The NBR model assumes the dependent variable (Y) as a function of count data (vaccination doses) following the negative binomial distribution. Unlike the regular regressions, the variance and mean do not need to be equal in the NBR model. The prediction of the dependent variable for an observation "i" is calculated using the following equation:

$$Pr(Y = y_i|\mu_i, \alpha) = \frac{\Gamma(y_i, \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)}\left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}}\left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i} \quad (11)$$

where μ is the mean rate of y per unit of time, α is the reciprocal of a scale factor, and Γ is the gamma distribution.

We first trained the model based on the current vaccination rates in all countries. Then for each of the countries considered in this study, the current vaccinates data (as of May 22nd, 2021) was ingested into the model and the future vaccination rates were simulated. Based on the simulation results, the approximate time at which the countries would reach herd immunity was estimated. Herd immunity was considered when the countries vaccinate 60% of their population with two doses of vaccine. We applied LSTM Model in Python using "TensorFlow" and "Keras" packages and the NBL model was applied using the "statsmodel" package. We highlight that the performed analysis is solely based on the currently available data (end of March 2021) and subject to bias due to the future changes in vaccination policies imposed by the countries.

3. Results

As of March 2021, the world is still struggling with the COVID-19 pandemic. During the past year (from March 2020 to March 2021), the epicenter of the COVID-19 in Asia has been changing from China to Iran, India, Russia, and Turkey. Currently (As of March 2021, India, with over 11.7 million confirmed cases and 160,000 deaths, has the highest COVID-19 reports among the Asian countries, followed by Russia with

4.4 million cases and 93,800 deaths (while Georgia has the largest COVID cases and mortality per capita). Among the eight countries considered in this study (shown in Fig. 3), United Arab Emirates has the highest cases per capita (42,500 per one million population), and Iran has the highest mortality rate (722 per one million population). With only 1850 COVID-19 cases and 32 deaths per capita, South Korea had the lowest COVID-19 statistics, depicted also in Fig. 3 (bottom rows). While several socio-cultural, economic, environmental factors interact with each other in each country for dealing with the COVID-19 pandemic, in the following sections, we have broken down how certain factors affect the transmission rate of COVID-19. Fig. 3 summarizes the statistical reports for COVID-19 over Asia and for the selected countries.

3.1. Correlation assessments

We used GLMs to investigate the correlation between the COVID-19 cases, mobility, air quality, air temperature, relative humidity, and governmental response. The resulting pair-wise correlations and statistics (based on the best fitted GLM) are summarized in Fig. 4 in a matrix format. The sub-figures on the principal diagonal of the matrix show each variable distribution. Overall, we only found mobility to have a significant impact on virus transmission. The next subsections discuss details of each variable role.

3.1.1. Meteorological variables

For air temperature (Tair) and relative humidity (Rh), the weekly relative changes in Tair and Rh were examined against the number of infected cases using the Gaussian model. As depicted by Fig. 4, there was no significant correlation between the weather-driven parameters and the number of COVID-19 patients. For the entire period from March to December 2020, the relationship between the Tair, Rh, and COVID-19 cases resulted in statistically non-significant correlations ($P_{value} \gg 0.05$). Therefore, the impact of meteorological factors on the virus transmission was found to be unlikely. There was no static upward or downward trend in the number of the COVID-19 infections with the increase or decrease in air temperature (or relative humidity) during different months, signifying the transmission rate is not driven by the change in the weather-driven parameters. The meteorological variables did not also correlate with other parameters considered in the study.

Thus, studies and analyses focusing on only a limited time window or specific geographical location could be largely biased. A number of early-stage studies (e.g., Gupta et al., 2020; Ma et al., 2020; Tosepu et al., 2020; Xie and Zhu, 2020) have reported specific ranges for air temperature and relative humidity at which the virus has its most efficiency. However, these studies do not consider the impact of various spatiotemporal scales in the analysis.

Overall, our analysis showed that meteorological variables do not affect virus transmission or prevention. Similarly, Jahangiri et al. (2020) analyzed the effect of absolute temperature (AT) on virus transmission in 31 different Iran provinces and reported on the low sensitivity of AT and the COVID-19 patient. Jamshidi et al. (2020) performed a similar analysis globally and reported no impact from the weather on the virus transmission. Therefore, there is no scientific reason to have fewer patients in warmer climates than in moderate or cold climates.

3.1.2. Mobility

The correlation between mobility and the number of COVID-19 cases was investigated using the Poisson model. It should be noted that mobility data used in the analysis was computed as the difference in the transportation during the pandemic with respect to a similar time during pre-pandemic (negative values show reduction and positive show increase in mobility). It was found that the change in mobility was strongly correlated with the COVID-19 pandemic growth change. While the resulting models fitted to the data were not adequately representative (given the high magnitude of $\chi^2 = 6623$), the resulting P_{value} (<0.01) shows the significance of the resulting correlation coefficients.

Our analysis depicted that the reduction of mobility was in concert

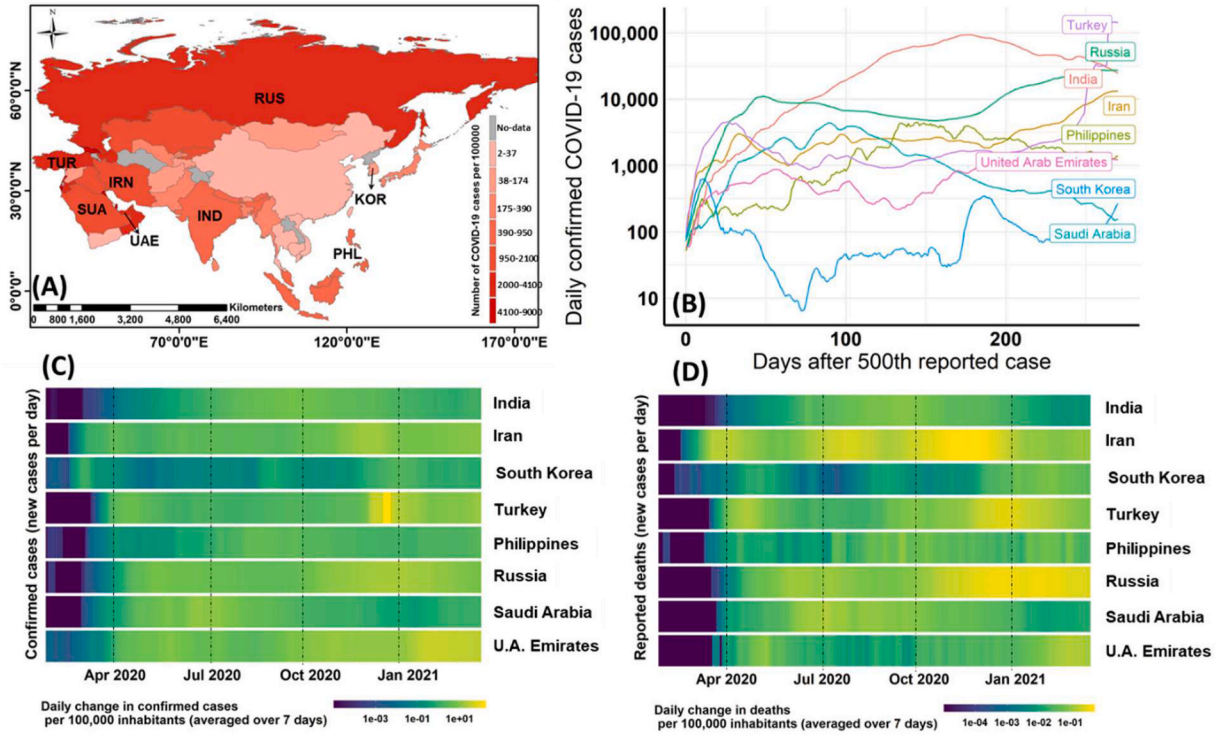


Fig. 3. Overall statistics of COVID-19 infected cases and mortality rate in Asia with specific attention to Iran (IRN), Turkey (TUR), Russia (RUS), India (IND), Saudi Arabia (SAU), United Arab Emirates (UAE), Philippines (PHL), and South Korea (KOR). The spatial distribution and time series of daily COVID-19 cases are shown in the top row, and the bottom row shows the changing rate of the daily COVID-19 cases and mortality from March 2020 to March 2021.

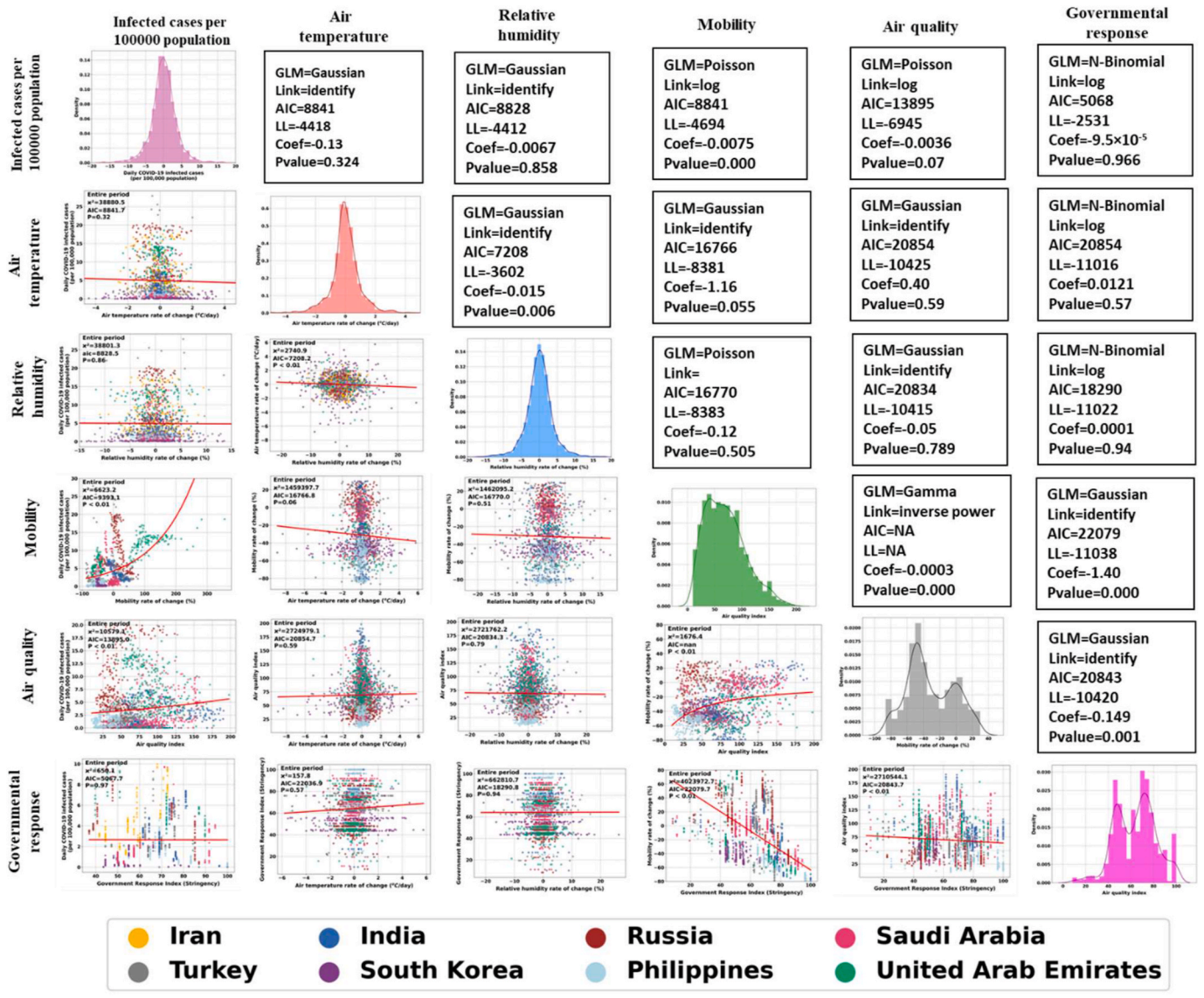


Fig. 4. The relationships between the COVID-19 infected cases and air quality, air temperature, relative humidity, mobility, and governmental response based on the best fitted GLM in Iran, Turkey, India, South Korea, Russia, Philippines, Saudi Arabia, and the United Arab Emirates (from March to December 2020). The sub-figures on the principal diagonal of the matrix show each variable distribution and upper diagonal elements show the statistics for each of the models.

with the dropping COVID-19 cases, and when there was a surge in mobility, COVID-19 transmission increased rapidly. Taking the United Arab Emirates as an example, we can see that when mobility increased from 50% to 200%, the number of infected cases per capita increased from 5 to 15. The link between daily changes in mobility and COVID-19 cases was observed during all months with different levels of strength for the countries investigated.

Although the relationships between the mobility and COVID-19 pandemic has not been much discussed at cross country level, it has been highlighted at the country or city level in a number of studies (Badr et al., 2020; Engle et al., 2020; Nouvellet et al., 2021). Mobility was also found to have a significant impact on the level of pollutants as shown in Fig. 4. A higher transportation rate would directly contribute to the level of particulate matter in the atmosphere and change air quality.

3.1.3. Governmental responses

To investigate the correlation between the stringency index and COVID-19 transmission rate, we applied the negative binomial model. The governmental response yielded a non-significant ($P_{\text{value}} = 0.966$) correlation with the COVID-19 cases, with overall x^2 of 650. The reason for the absense of correlation between the GSI and COVID-19 was the consideration of all eight countries at once. As depicted in Fig. 5 (B and C), there were two distinct patterns in the resulting GSI. Exploring the two opposite patterns together resulted in the non-significant pattern shown in Fig. 5, A.

Expecting a lower number of COVID-19 cases with a higher stringency index would be a logical and reasonable hypothesis and was observed in Russia, United Arab Emirates, India, and the Philippines (Fig. 5, B). Nevertheless, a negative trend (increased number of COVID-19 cases with higher GSI) was detected for certain countries. The negative correlation (upward trend) likely signifies a delayed governmental response in imposing mitigating strategies against the COVID-19 pandemic. Iran, Turkey, and Saudi Arabia were the examples where the governmental response showed a positive correlation with the COVID-19 transmission rate (Fig. 5, C).

Taking the case of Iran as an example, the country publicly reported the first 43 COVID-19 cases in mid-February, and with no strict quarantine or public health measures, the country reported more than 29,000 cases by mid-March 2020 (Analysis of Traveler Data Suggests Iran's COVID-19 Trajectory Far Higher Than Officially Reported 2020.

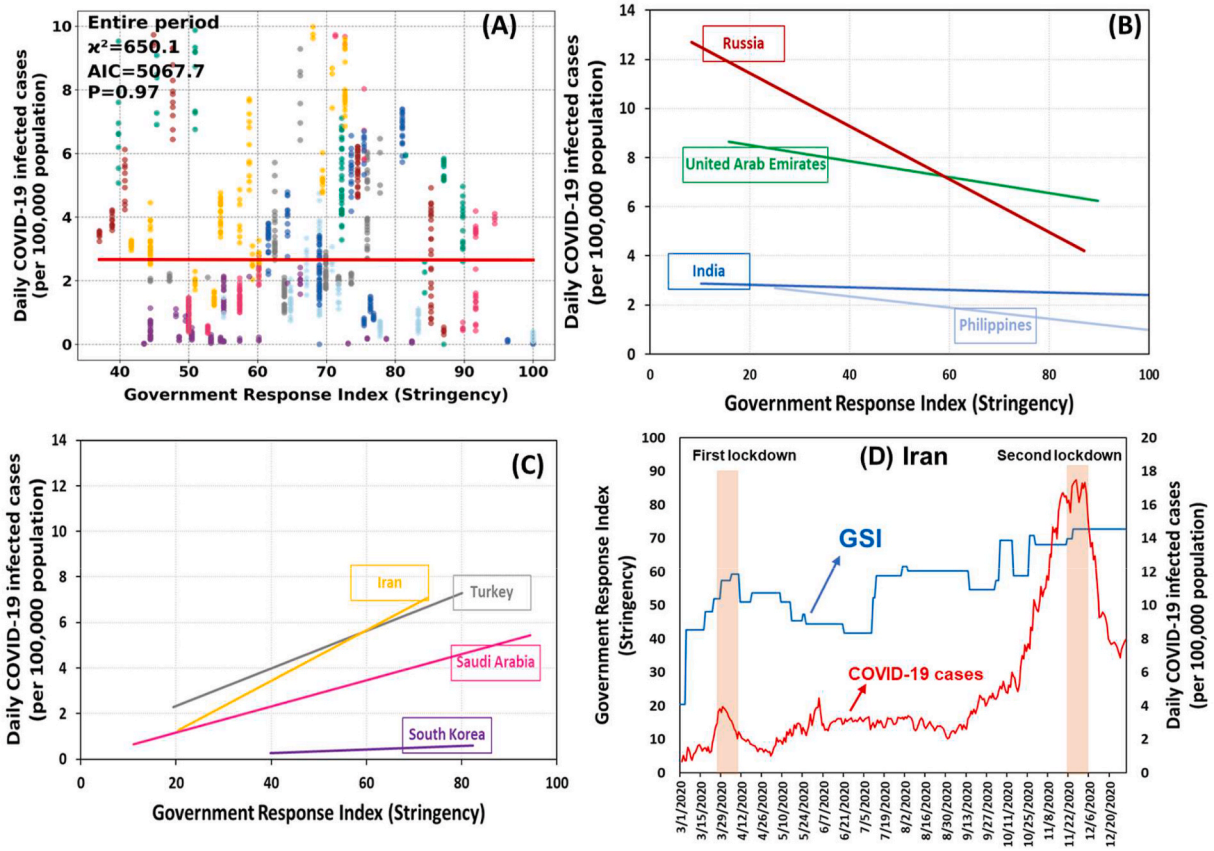


Fig. 5. The relationships between the COVID-19 infected cases and the governmental response index (GSI) in Iran, Turkey, India, South Korea, Russia, Philippines, Saudi Arabia, and United Arab Emirates (A) and the overall trend in each country (B and C). An example of time series of the GSI and daily COVID-19 cases and lockdown periods is shown in subfigure D. The data period is from March to December 2020.

Available from: https://www.contagionlive.com/view/analysis-of-traveler-data-suggests-irans-covid19-trajectory-far-higher-than-officially-reported/). After a sharp increase in the COVID-19 cases in March, Iran has imposed a national lockdown from March 20th to April 3rd, 2020. Another more strict national lockdown was imposed from November 21st to December 4th, 2020, again after a radical surge in the number of cases (from less than 5000 daily cases on October 21st to more than14000 daily cases on November 21st). Although these policies were successful in terms of curbing the pandemic growth (reduced nearly 40–50% of daily cases), they were imposed late and after the virus surge, causing the negative trend depicted in Fig. 5, D.

Among the studied countries, South Korea had the best performance in governmental responses. As demonstrated in Fig. 5-C, with the onset of the COVID-19 pandemic, the country imposes policies with a high GSI (>40), nearly twice stricter than other countries in our study. The rapid and strict governmental response in South Korea was further increased during the pandemic and provided an excellent example of curbing the virus transmission. Therefore, the early response would be a key for controlling an outbreak like COVID-19.

It should be highlighted that while China was not among our studied countries, its implementation of multifaceted public health measures and aggressive non-pharmaceutical interventions to control the COVID-19 pandemic can be seen as an excellent example for future efforts to combat global health emergencies. Intensive intracity and intercity traffic restriction, social distancing measures, cancellation of social gatherings, and home quarantine were among the main implemented interventions (Leung et al., 2020). Several studies have shown that the effective reproduction number (Rt) gradually decreased in Wuhan, China (as the first epicenter of COVID-19 pandemic) in late January 2020, when the government blocked all outbound transportation from the city, suspended public transit within the city, and implemented strategies such as compulsory mask-wearing in public places (Pan et al., 2020). The Rt further decreased by improvement in medical resources, the government implemented a stay-at-home policy in the city and initiation of door-to-door and individual-to-individual symptom screening for all residents (Hao et al., 2020; Pan et al., 2020).

3.1.4. Air quality

As discussed in the material and method section, we used satellite-driven AOD data (for the short-term analysis) and PM2.5 (for the long-term analysis) for the air quality metric. The correlation between the AOD and COVID-19 infection during the study period was investigated using the Poisson model. Based on our analysis n the short term, while we found footprints of higher infectious rates in more polluted areas, the correlation between the two was not statistically significant ($P_{value} = 0.07$, x^2 of 10,579). We highlight that exploring the role of air quality on the COVID-19 health outcome using the regression analysis is challenging because of the chain effect that exists between the variables in an urban setting. The sources of air pollution are intrinsically linked with transport and industry. The COVID-19 pandemic's direct impact on shutting down transportation and business would translate into air quality changes. This interconnectivity between the variables indicates the importance of the causal analysis presented in the following subsection.

While a number of studies have shown the link between the air quality and COVID-19 mortality rate using the data during the pandemic (Fattorini and Regoli, 2020; Wu et al., 2020), we emphasize that comparing the number of COVID-19 mortality and the air quality data would not be the optimal approach to identify the associations between the two variables. Analyzing such attribution requires the exposure data

at a personal level. In a recent work by Wu et al. (2020) using a long-term analysis for capturing the impact of air pollutants on the COVID-19 mortality rates was suggested. They have analyzed a 17-year average of PM2.5 concentrations (using satellite data) over the United States and reported that for every 1 μg/m^3 increase in air pollution, there would be an 11% increase in mortality from COVID-19 infection. In a similar manner, we have used the PM2.5 data over an 18-year period (1998–2017) to check the long-term exposure health outcome concerning COVID-19. Fig. 6 shows the spatial distribution of PM2.5 and COVID-19 deaths over the Asian countries. The cross-country correlation (scatter plot in Fig. 6) did not show any significant correlation between the air quality and the higher risk of COVID-19. We caution that evaluating the association between air pollution and COVID-19 transmission or mortality rates at the countrywide level may not be an optimal approach. This is particularly due to each country's specific socio-economic situation and the measures that have been taken during the pandemic that may interfere with the outcome. For example, in the eastern part of China, the air quality is extremely low (PM2.5 $>$ 50 μg/m^3); however, the policies and measures that China has imposed during the pandemic have led to a low mortality rate (1–2 per 100,000 inhabitants).

Therefore, we further analyzed the long-term exposure taking the examples of Iran, India, and South Korea. These countries were selected because we had access to state-level COVID-19 data. As demonstrated in Fig. 7, the more polluted areas were associated with a higher risk of severe COVID-19 outcomes. The hotspots of air pollution in Iran are located in the southwest, west, and northwest of the country (up to 27 μg/m^3), where we found a higher number of COVID-19 death. In particular, the Khuzestan province (in southwest Iran) has been a hotspot for COVID-19 deaths with nearly 200 per 100,000 population. This region has been suffering from low air quality due to frequent dust storms (Goudarzi et al., 2019), which could have contributed to the susceptibility of the inhabitants to respiratory disease, including but not limited to COVID-19. We should highlight that air quality could be one of the underlying factors for a higher COVID-19 death rate in Khuzestan province, as other factors (in particular socio-cultural factors) would be important as well. For example, excessive social gatherings and local ceremonies in the Khuzestan province (based on the authors' local knowledge) could cause a higher number of COVID-19 cases and a higher likelihood of COVID-19 death in this province.

In India, the center to the northern part of the country experiences one of the worst air quality (PM2.5 $>$ 60 μg/m^3), and in these areas, the likelihood of COVID-19 mortality was higher. However, in southern India, despite a relatively better air quality compared to the central and northern parts, the COVID-19 death rate was higher. This inconsistency could stem from the coarse spatial resolution COVID-19 data (although we have scale it down to state-level). Such complex problems require fine resolution data at a personal level, as other variables such as policies, urbanization, demographic information, and the healthcare system could also possess important roles in changing the COVID-19 transmission rate (Jamshidi et al., 2020; Mallires et al., 2019). For example, southern states in India have less developed health infrastructure, particularly in rural areas, compared to the northern states (Kumar et al., 2020). The inefficient healthcare system in these states may contribute to the higher number of mortalities despite their better air quality. Maharashtra state, located in southwestern India, is the second-most populous state and one of the most urbanized states (more than 50%) in India. These factors could cause rapid COVID-19 transmission in this state. The high infection rate and failure of the health system during the second wave of COVID-19 likely caused the highest mortality rate in Maharashtra (After Delhi) as shown in Fig. 7. These examples clarify the impact of other factors driving the COVID-19 mortality rate and the low correlation with air quality.

In South Korea, the western and northwestern regions suffer from higher particulate matter and higher mortality rates. However, the positive correlation between the pollutants and COVID-19 mortality rate, unlike Iran and India, was not statistically significant. One reason for this inconsistency could be the intermediate factor of urban density, which highly contrasts with the administrative boundaries of South Korea. The role of urban density is also noticeable in the case of Iran and India (scatter plots, Fig. 7), where higher PM2.5 values and mortality rates were found in denser areas (bigger circles in the scatter plots). The administrative divisions of South Korea (for which the COVID-19 data was available) consists of heavily dense areas like Seoul (16,150

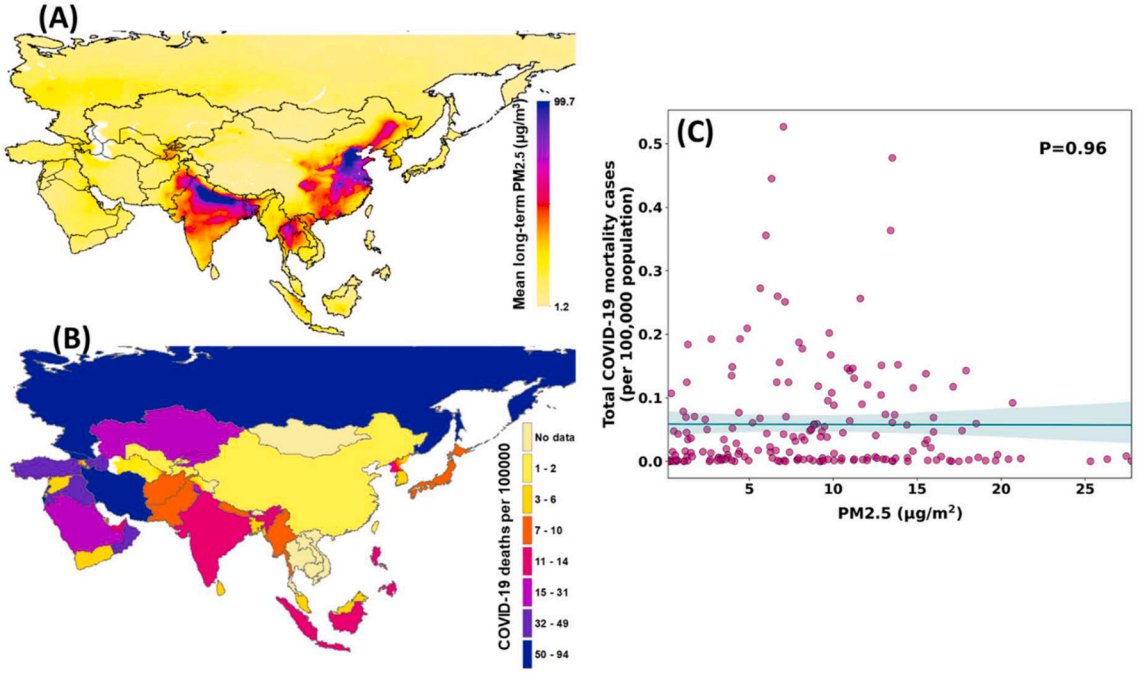


Fig. 6. The spatial distribution of the mean long-term PM2.5 pollutants from 1998 to 2017 (A), total number of COVID-19 deaths (from March to December 2020) (B) across the Asian countries. The correlation between A and B is shown in subfigure C for all the nations with available data.

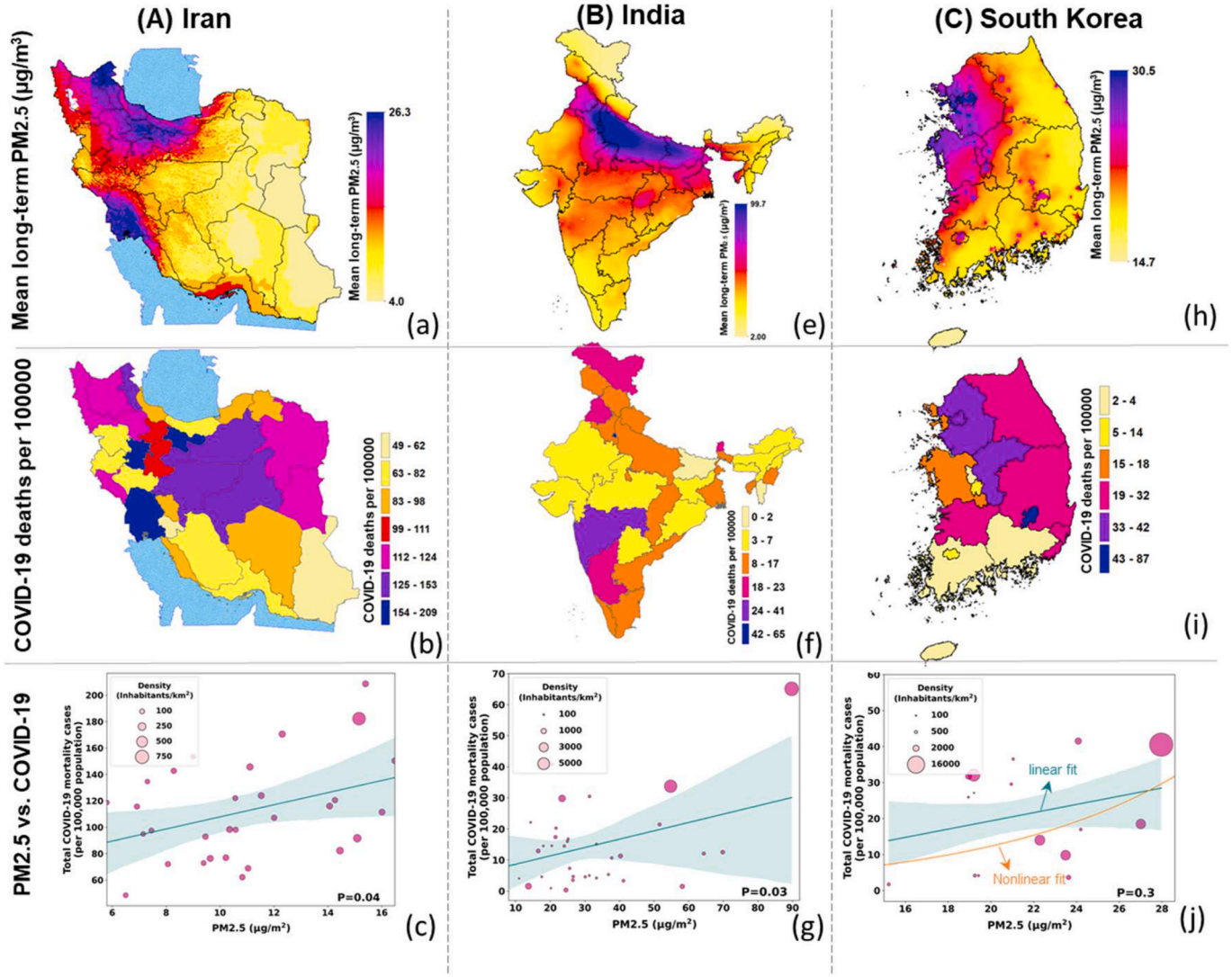


Fig. 7. The spatial distribution of the mean long-term PM2.5 pollutants from 1998 to 2017, the total number of COVID-19 deaths (from March to December 2020) for Iran (column A), India (column B), and South Korea (column C). The correlations at the state (or province) level are shown in the scatter plots (the size of circles show the density of the states and the lines are fitted using generalized linear models).

inhabitants/km^2) or Ulsan (4450 inhabitants/km^2), and less dense regions like Gangwon (75 inhabitants/km^2) or Gyeongsangbuk (144 inhabitants/km^2). As urban density is directly linked with air quality, mixing the areas with different urban densities could lead to weak correlations (McCarty and Kaza, 2015). As shown in Supplementary Figure S1, we divided the states in South Korea into low-density areas (<500 inhabitants/km^2) and high-density regions (>500 inhabitants/km^2) and found a stronger correlation between air quality and COVID-19 mortality in less dense areas. The interplay between the pre-existing intermediate factors (e.g., socio-cultural or demographic factors that are explained above) becomes more complex in high-density areas. In these areas capturing the correlation between air quality and respiratory diseases are more challenging and cannot be done by regression analysis. Another intermediate factor affecting the PM2.5-mortality correlation in South Korea could be its strict public health strategies. South Korea instituted and maintained effective and immediate containment and mitigation strategies, including widespread, free and rapid testing, meticulous tracing of all contacts, and crowd avoidance, which could have greatly impacted the associations between COVID-19 and environmental variables, including air pollution.

Although our analysis showed a potential link between long-term exposure and COVID-19 health outcome, we caution that these analyses only provided an overall insight as we did not consider other cofounding parameters in the long-term analysis that could potentially interfere with the results. Nevertheless, the insights provided by our study create an avenue for conducting a more detailed analysis considering other environmental and socio-economic factors.

3.2. Causal analysis

We also performed the causality analysis for each of the countries (separately) and all together. Our results are shown in Fig. 8 (full results with all countries are included in the Supplementary Fig. S2). As discussed in the Material and method section, ten time-lagged steps were analyzed (i.e., the relationships between the variables at each time steps were analyzed with respect to the last ten days $\{i_{-10}, i_{-9}, \ldots, i_{current}\}$. Therefore, Fig. 8 includes four key features: the auto-MCI, cross-MSI, lag time, and correlation arrows. The auto-MCI shows the correlation for each variable with itself at previous time steps and displayed using the colored circle around the variable. The cross-MCI signifies the correlation between the variables and is shown with the colored arrows. The blue-colored arrows indicate a negative correlation (ranging from −1 to 0), and orange arrows show a positive correlation (ranging from 0 to 1).

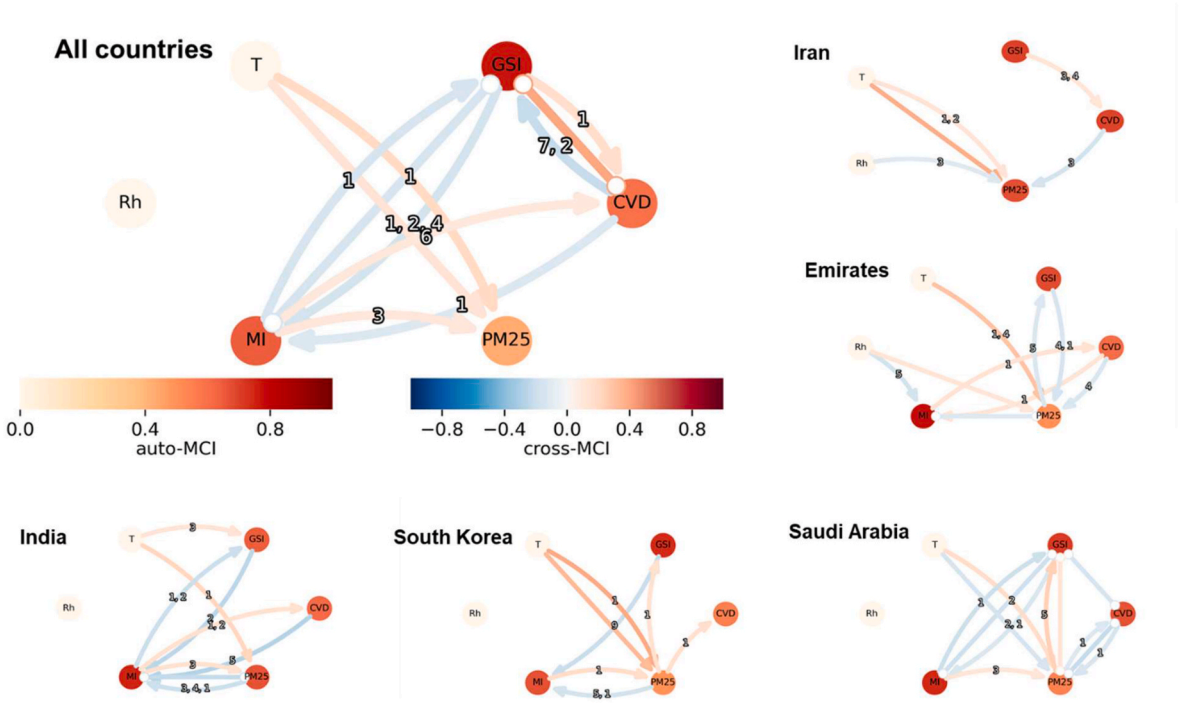


Fig. 8. The causal dependencies of air temperature (T), relative humidity (Rh), Mobility (MI), air quality (PM25), governmental response (GSI), and the number of COVID-19 infections (CVD). The arrows show the cross-variable MCI, with the head showing the parent and tail showing the child. The numbers on the arrows show the lagged-time step. The circle around each variable shows the self-dependency (Auto-MCI). The data is based on the March to December 2020 period, and examples of the individual countries are shown as subfigures.

The numbers on the arrows show the time step at which the correlation occurred.

While different countries demonstrated a slightly different causality pattern among the variables, the overall concept was similar. Accordingly, the connection between the number of COVID cases and mobility was found to be mutual. A change in mobility has been translated to a shift in the COVID-19 cases with a six day-time-lag (on average). That means based on the data from March to December 2020, whenever the mobility was reduced, its positive impact on curbing the COVID-19 transmission could be observed within six days. On the other hand, a change in COVID-19 has also triggered a reverse change in mobility with a 3-day-lag-time. Considering COVID-19 as a driver of mobility caused the reverse correlation and likely occurred due to a 'fear concept'. More specifically, the higher COVID-19 cases were caused people to be more cautious, and therefore, the number of daily trips has been reduced, shaping the reverse correlation.

The governmental response was linked to the COVID-19 cases and mobility. As shown in Fig. 8, GSI impacted COVID-19 cases with a positive link (given GSI as the parent) and a negative link (given COVID-19 cases as the parent). The reversed links conveyed that (on average) the imposed policies, while effective, were not timely mannered. The reversed links between the GSI and the number of infected cases confirmed that stricter strategies (higher GSI) were imposed with a rise in the COVID-19 cases. The GSI also caused mobility changes (negative cross-MCI), and the mobility changes have altered the virus transmission (as stated above). The links between GSI and mobility were negative, conveying that higher GSI caused lower mobility levels with a 1–4 days' lag.

No links were found between the PM2.5 with the number of COVID-19 cases. There were also some links between mobility and air temperature to PM2.5, which were reasonable. A number of studies have demonstrated the connection between the meteorological variables and the air quality variations (Bloomer et al., 2009; Chen et al., 2019; Koken et al., 2003). The positive link between higher mobility and higher air pollution was expected as transportation has consistently been recognized as a significant air pollution source. These links are crucial for a factual analysis of the factors affecting COVID-19 cases. In particular, the climate condition may not directly affect the COVID-19, but it may cause a change to other variables (for example, air quality) and thus, indirectly affect the COVID-19 condition.

3.3. COVID-19 exit strategies to normal condition

This section presents the VirSim model simulation results considering different strategies for a transition from the COVID-19 pandemic to a normal condition. As discussed in the material and method section, the simulations were carried out for a country with ten sub-regions (states or provinces), each comprised of equal divisions (e.g., 100 divisions) and a population of 1000 people living in each division. Fig. 9 shows how the execution of each strategy affects the number of COVID-19 cases over time. The horizontal line at 6000 cases per million reflects the healthcare capacity (the value was originally developed by the model for Netherland, but here since we are comparing different countries with different healthcare capacities, we leave 6000 cases per million as a plausible consideration). The vertical dashed lines show the moment at which intervention occurs. Each strategy's execution affects the percentage of the population recovering from the infection over time in a different manner which is visualized in the Supplementary Fig. S3. The horizontal dashed line at 60% indicates the minimum level of herd immunity required to avoid an epidemic (de Vlas and Coffeng, 2021).

In the radical opening, as shown in Fig. 9, A, the number of infected cases dramatically increases over a short period of time, exceedingly far beyond the healthcare capacity. The radical opening strategy issues a rapid exit from the pandemic with achieving herd immunity within the first 100 days (Fig. 9, A) and has economic justifications for a country (del Rio-Chanona et al., 2020). Nevertheless, it causes a high number of mortality cases.

Considering the Phased Lift of Control (PLoC) strategy, we have evaluated four variations, including 10 and 20 lifting phases with 30 and 60 interval days (Fig. 9, B, and C). Lifting the imposed policies at 10

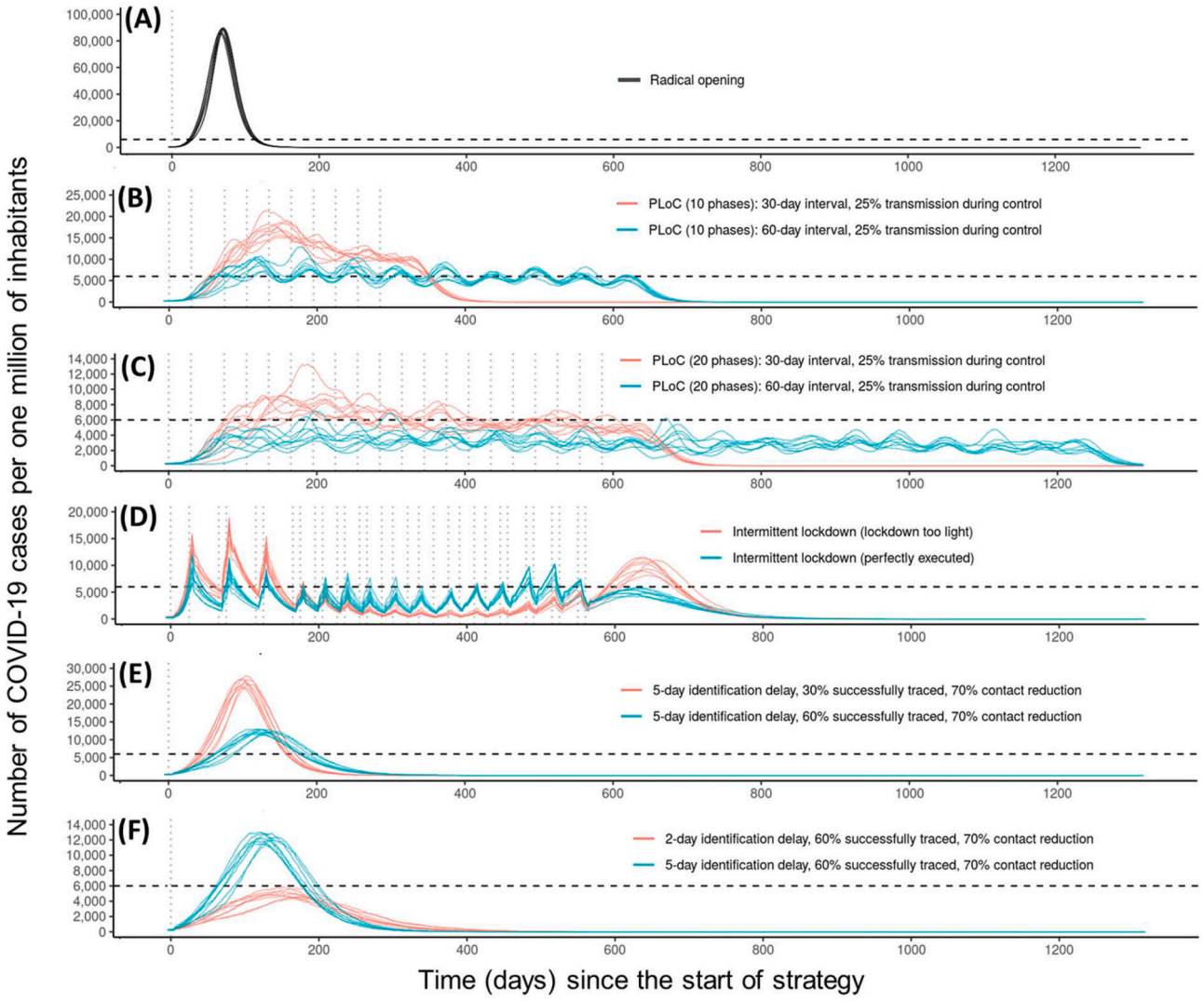


Fig. 9. The simulated number of COVID-19 cases resulting from the VirSim model (de Vlas and Coffeng, 2021) for exiting the COVID-19 pandemic based on radical opening (A), Phased Lift of Control (PLoC) with 10 phases (B), and 20 phases (C), intermittent lockdown (D), contact tracing with different tracing percentage (E) and different identification delays (F). The vertical dashed lines show the moments at which an intervention was removed, and the horizontal dashed line shows the healthcare capacity (fixed at 6000 beds per one million). The two lines (red and green) show different settings used in the exit strategy (also showed in the legend). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

phases with 30-day intervals overburdens the healthcare capacity with a number of cases fluctuating from 10,000 to 15,000 per million; however, the herd immunity achieves with 10 months (red lines in Fig. 9, B) and ends the pandemic condition entirely in less than 15 months. Increasing lifting intervals of the lockdowns to two months (60-day intervals) would end the pandemic in nearly two years (achieving the herd immunity in ~17 months, green lines in Fig. 9, B), but the healthcare should be operating at its full capacity during the two years. For countries with a lower healthcare capacity, like Iran with 1600 and the Philippines with 1000 hospital beds per one million populations (according to the world bank data (2021). www.data.worldbank.org/indicator/SH.MED.BEDS.ZS, and National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/nuccore), the 10-phased lift strategy has a huge cost for the healthcare system. In contrast, countries like South Korea, with 12,000 hospital beds per population, could better manage the 10-phase lift strategy. Shifting from 10 to 20 phases reduces the healthcare burden but significantly delays herd immunity and increases the pandemic duration. A 20-phase lifting strategy with 60-day intervals causes the number of infections to maximize at nearly 3000 cases per million (half of the healthcare capacity) but takes nearly four years to leave the pandemic condition entirely.

The intermittent lockdown strategy (Fig. 9, D) leads to a spike in COVID-19 cases after each lockdown is ended. When a lockdown is executed sub-optimally (light lockdown), the transmission rate was considered 30%, and in a perfectly executed lockdown, the transmission rate was 25%. With a light lockdown (red lines), the number of peaks is twice higher than the healthcare system's capacity for the first six months and then decreases. Compared to a light lockdown, a perfectly executed lockdown leads to a more controlled virus transmission with smaller waves during the first six months but slightly higher peaks during the pandemic, which took nearly 26 months to end. The intermittent lockdown is not an ideal strategy as it creates a critical moment overloading the healthcare system (with 6000 hospital beds per million). Both intermittent lockdown strategies bring herd immunity at nearly two years (Fig. 9, D).

The last two rows of Fig. 9 (E and F) show the contact tracing strategy simulations. The policy was found to be very sensitive to the delayed time for identifying the host and the rate of success in tracing the infected cases. Considering a 5-day delay in tracing the infected cases

and reducing their contact with other individuals at 70% level, there was a large difference between 30% and 60% tracing levels. When only 30% of the infected cases were identified, there would be a large surge in the COVID-19 cases because 70% of the infected cases could freely interact with others and transmit the virus (the red lines in Fig. 9, E). Doubling the success rate of tracing the virus hosts to 60% would flatten the curve and reduces the number of infected people in the community (the green lines in Fig. 9, E). Nevertheless, a 5-day delay in tracing will not be optimal for the healthcare capacity. Reducing the delayed time from 5-day (the green lines in Fig. 9, F) to 2 days (the green lines in Fig. 9, F), the maximum number of infected cases would not surpass the healthcare capacity with peaking at nearly 6000 patients for a short period of time. The contact tracing variations were not sensitive to time as they showed one year to get back to a pre-pandemic condition (i.e., normal condition). In non-of the variations of the contact tracing strategies, the herd immunity would be achieved (Fig. 9E and F) as this strategy boils down at the individual level and puts the lockdown conditions on individuals, not the community. Therefore, there would be a limited community spread and no herd immunity.

Among all the strategies, contact tracing seems to work the best if executed correctly. With a 2-day delay in tracing the virus hosts, a 60% success rate of tracing the hosts, and a 70% contact reduction with the hosts, a pandemic will end in a year without overburdening a healthcare system with 6000 capacities per million. Note that all these simulations were carried out considering a 25% transmission rate without considering vaccination. Vaccination could significantly impact the simulated strategies. It should be noted that while we did not perform an in-depth risk assessment for each strategy, our visualizations partially cover this shortcoming. Considering only one average healthcare capacity for all countries and simulations, we can consider the risk probability of each scenario exceeding the healthcare capacity close to one (when the simulated infected cases break the healthcare capacity line) or close to zero (for the simulated curves below the healthcare threshold).

The factors evaluated in this study were only examples that could potentially affect the virus transmission. Other factors could have been triggered for the analysis, yet we skipped those factors due to the text limitations and data availability. For example, the virus mutation could affect the transmission rate, and because the SARS-CoV-2 is an RNA virus, frequent mutations are expected (Holland et al., 1982). One example of such mutations that become the dominant mutation in Iran is D614G. This mutation occurs in spike protein leads to an amino acid change from aspartate to glycine at position 614 (D614G) (Fattahi et al., 2020). As studies suggested the viral isolates which carry D614G mutation have a significant role in increased infectivity and transmission by enhancing viral loads in the upper respiratory tract of COVID-19 patients (Korber et al., 2020; Plante et al., 2020; van Dorp et al., 2020; Volz et al., 2020).

Another factor could be mask usage among the community. The positive role of mask-wearing in reducing the droplets and air dispersion from a coughing incident has been well-established in clinical studies (Tang et al., 2009). Non-clinical studies have also shown that countries such as Hong Kong, South Korea, Vietnam, Singapore, and China, which started using face masks earlier, could control the disease spread and return to pre-COVID condition more efficiently. Based on these studies, people's habit of mask usage was already widespread before the COVID-19 pandemic (i.e., generally to protect people from respiratory viruses such as MERS, SARS, and seasonal flu or as a reaction to air pollution) in such countries (Han et al., 2020; Nishiyama et al., 2008) (MarketWatch website. 2020. https://www.marketwatch.com/story/what-we-can-learn-from-south-koreaand-singapores-efforts-to-stop-coronavirus-in-addition-towearing-face-masks-2020-03-31). The data on the mask use is limited and mostly based on non-official surveys, which may not sufficiently represent the spatial distributions of the mask usage among the individuals, and thus, was not applicable in our study.

3.4. COVID-19 vaccination prediction

One year after the beginning of the COVID-19 outbreak, countries worldwide are racing to provide and distribute vaccines for their citizens. As of May 2021, over 1.7 billion COVID-19 vaccine doses have been administered worldwide, out of which approximately 52% (893 million) have been injected in Asia. Vaccination plans are still in the early stages in most Asian countries and differ from country to country. The vaccine rollouts in Asia are affected by factors such as the economic potential of countries, vaccine skepticism/mistrust among the citizens, and an overcautious attitude towards the vaccines (Cable News Network (CNN), 2021, https://www.cnn.com/2021/01/08/asia/asia-covid-vaccines-explainer-intl-hnk-scli-dst/index.html). In this section, we focused on the vaccination data for the eight countries in Asia.

As of May 2021, India, with around 194 million doses, has the largest number of administered doses of vaccines, while Iran and the Philippines have the least vaccinations. The lagging vaccination rate in the Philippines has been partially attributed to the public anxiety stemming from the association between the Denvaxia vaccine and children's deaths (Cohen, 2019). In Iran, the bans against western-manufactured vaccines have been reported as the primary reason for the slow vaccination rate (available at: https://www.theguardian.com/world/2021/jan/08/untrustworthy-covid-vaccines-from-us-and-uk-banned-by-iran). While India outperformed the other seven countries with respect to the number of vaccines, UAE has vaccinated the highest ratio of its population.

The vaccination simulation results based on the deep learning approach and using the available vaccination data (from December 2020 to May 2021) are presented in Fig. 10. The simulation includes an uncertainty range (shown as a red shaded area in Fig. 10 with upper and lower bounds). The upper and lower bounds were defined based on the weight matrices and bias vector parameters that were primarily influenced by the data in countries with high and low vaccinate rates, respectively. Accordingly, the upper bound shows an optimistic scenario with the maximum vaccination rate, and the lower bound defines a pessimistic scenario with the minimum vaccination rate.

Based on the simulations, after UAE (which has already reached the herd immunity stage), Saudi Arabia would be the first country that inoculates its population (i.e., 60% fully vaccinated population) around August to September 2020, significantly faster than other evaluated countries. Turkey and South Korea could be the next countries to reach herd immunity ambitiously by the end of 2021 (which could extend to March 2022). Russia and the Philippines showed a higher range of uncertainties due to the various vaccination rates administered in these countries with the onset of vaccination. Accordingly, in an optimistic scenario, Russia and the Philippines would reach herd immunity by early 2022. However, in a pessimistic scenario (based on the lower bound simulations), reaching vaccination-based herd immunity could be postponed to the end of July in the Philippines and October in Russia. India and Iran would be the last countries that vaccinate their people (October to November 2022). For India, the high population and limited vaccine supply (due to the sharp rise of COVID-19 in February-April, 2021 and impeded vaccine development) would be the primary causes of the delayed herd immunity. For Iran, the late onset of vaccination and slow vaccination rates would likely contribute to the delayed herd immunity. It should be noted that our simulations hypothesize that no other catalyst or inhibitor factors interfere with the future vaccination rate (which can be violated given real-world data dynamics and complex behaviors). For example, Iran has been developing its own vaccine which could help to achieve a faster than predicted herd immunity. Another complex intervention is the delayed COVID-19 vaccine supply and export by India. India is the home of the world's biggest producer of COVID-19, nevertheless, with the devastating second wave of COVID-19 infection in the country during February-March 2021, the country has halted vaccine exports to focus on domestic inoculations (Gettleman et al., 2021). The vaccine shortage by India has also slowed down the

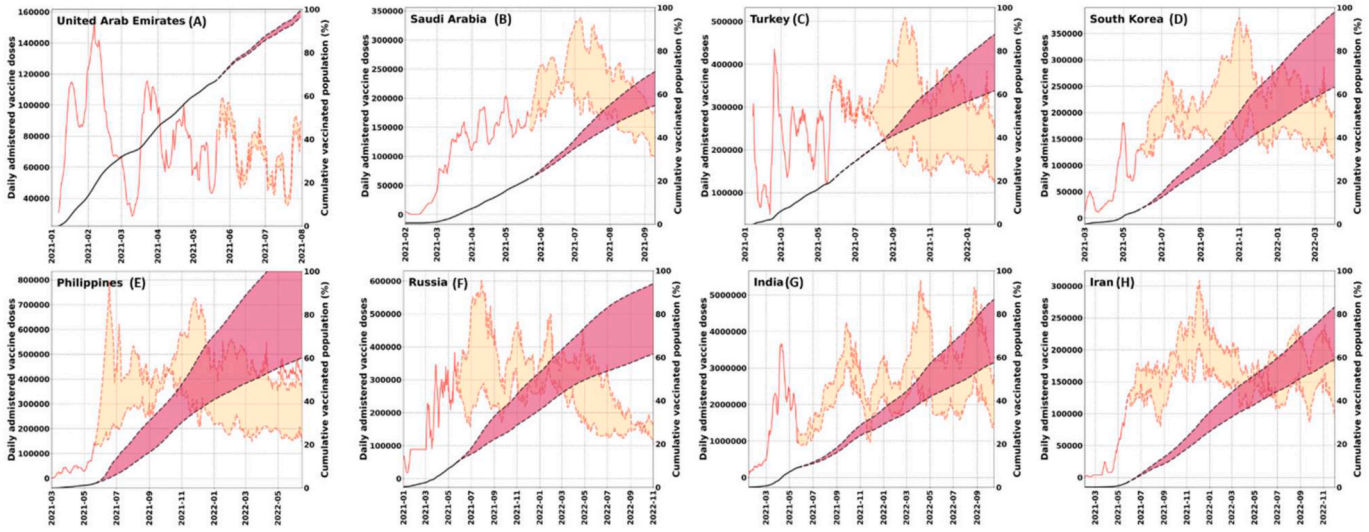


Fig. 10. The simulated daily administer COVID-19 vaccine doses (red lines) and cumulative vaccinated population (black lines) using LSTM deep learning approach based on the available COVID-19 vaccination data from December 2020 to May 2021. The solid lines show the real data, dotted lines show the predictions, and shaded areas show an uncertainty range based on optimistic and pessimistic scenarios. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

vaccination rates in countries that have ordered their vaccine supply from India (e.g., Iran).

4. Conclusion

Analyzing the real-world data specifically for an epidemic disease with complex dynamic interactions in day-to-day life is challenging and a bias-free analysis would not be possible without accurate high spatiotemporal resolution data. Nevertheless, we attempted to address the role of several parameters in driving the COVID-19 pandemic and evaluate different strategies that countries can adopt to exit the pandemic condition without overburdening the healthcare capacity. Our analysis is carried out at a countrywide scale to assess if the interactions between the environmental and socio-economic factors could be captured. Among the evaluated variables, we found mobility to be most effective for controlling the virus spread. Weather-driven parameters were found to be ineffective in driving the COVID-19 cases. A long-term analysis of air pollution and COVID-19 mortality revealed a higher risk of mortality for the population consistently exposed to low air quality. We highlight that an accurate estimate of air pollution's impact on COVID-19 cases (and other respiratory diseases) requires information at an individual level.

We also showed the efficiency of different policies for exiting the pandemic condition. While contact tracing was found to be more effective, selecting the best policies for each country depends on several factors, such as the healthcare capacity and the country's economic situation. All the tested policies were sensitive to the timing and duration of the interventions, and thus, executing them in practice requires diligent attention to the details discussed in the result section. Nevertheless, vaccine-induced herd immunity is the only long-term solution to protect the population from SARS-CoV-2 infection.

Our evaluations on the current vaccination status highlighted the socio-economic factors affecting the inoculation rates. The vaccine simulations differed significantly among the countries as UAE has already reached herd immunity (May 2021) and Saudi Arabia could reach this level by August to September 2021. Nevertheless, in countries like India and Iran, vaccine-based herd immunity could be achieved by late 2022. We should highlight that our simulations hypothesize that no other factor interferes with the current vaccination rates and the assumption can be violated given real-world data dynamics and complex behaviors.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2021.111657.

CRediT author statement

Maryam Baniasad: Conceptualization, Investigation, Formal analysis, Writing - original draft, Visualization. Morvarid Golrokh Mofrad: Conceptualization, Investigation, Writing - review & editing. Bahare Bahmanabadi: Conceptualization, Investigation, Writing - review & editing. Sajad Jamshidi: Conceptualization, Investigation, Formal analysis, Visualization, Supervision, Writing - original draft, Writing - review & editing.

References

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 modelling study. Lancet Infect. Dis. 20, 1247–1254.
Bloomer, B.J., Stehr, J.W., Piety, C.A., Salawitch, R.J., Dickerson, R.R., 2009. Observed relationships of ozone air pollution with temperature and emissions. Geophys. Res. Lett. 36.
Bubar, K.M., Reinholt, K., Kissler, S.M., Lipsitch, M., Cobey, S., Grad, Y.H., Larremore, D. B., 2021. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. Science 371, 916–921.
Cameron, A.C., Trivedi, P.K., 2013. Regression Analysis of Count Data. Cambridge university press.
Chen, C., Liu, X., Wang, X., Li, W., Qu, W., Dong, L., Li, X., Rui, Z., Yang, X., 2019. Risk of temperature, humidity and concentrations of air pollutants on the hospitalization of AECOPD. PLoS oone 14, e0225307.
Cheng, C., Barceló, J., Hartnett, A.S., Kubinec, R., Messerschmidt, L., 2020a. COVID-19 government response event dataset (CoronaNet v. 1.0). Nature Human Behav. 4, 756–768.
Cheng, Z., Lu, Y., Cao, Q., Qin, L., Pan, Z., Yan, F., Yang, W., 2020b. Clinical features and chest CT manifestations of coronavirus disease 2019 (COVID-19) in a single-center study in Shanghai, China. Am. J. Roentgenol. 215, 121–126.
Cohen, J., 2019. Controversy over Dengue Vaccine Risk. American Association for the Advancement of Science.

Conticini, E., Frediani, B., Caro, D., 2020. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environ. Pollut. 114465.
Damette, O., Goutte, S., 2020. Weather, Pollution and Covid-19 Spread: a Time Series and Wavelet Reassessment.
de Vlas, S.J., Coffeng, L.E., 2021. Achieving herd immunity against COVID-19 at the country level by the exit strategy of a phased lift of control. Sci. Rep. 11, 1–7.
del Rio-Chanona, R.M., Mealy, P., Pichler, A., Lafond, F., Farmer, J.D., 2020. Supply and demand shocks in the COVID-19 pandemic: an industry and occupation perspective. Oxf. Rev. Econ. Pol. 36, S94–S137.
Engle, S., Stromme, J., Zhou, A., 2020. Staying at Home: Mobility Effects of Covid-19 (Available at: SSRN).
Fattahi, Z., Mohseni, M., Jalalvand, K., Moghadam, F.A., Ghaziasadi, A., Keshavarzi, F., Yavarian, J., Jafarpour, A., elham Mortazavi, S., Ghodratpour, F., 2020. Two Independent Introductions of SARS-CoV-2 into the Iranian Outbreak. medRxiv.
Fattorini, D., Regoli, F., 2020. Role of the chronic air pollution levels in the Covid-19 outbreak risk in Italy. Environ. Pollut. 264, 114732.
Frontera, A., Cianfanelli, L., Vlachos, K., Landoni, G., Cremona, G., 2020. Severe air pollution links to higher mortality in COVID-19 patients: the "double-hit" hypothesis. J. Infect. 81, 255–259.
Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30, 5419–5454.
Gettleman, J., Schmall, E., Mashal, M., 2021. India Cuts Back on Vaccine Exports as Infections Surge at Home. The New York Times. https://www.nytimes.com/2021/03/25/world/asia/india-covid-vaccine-astrazeneca.html.
Goudarzi, G., Alavi, N., Geravandi, S., Yari, A.R., Alamdari, F.A., Dobaradaran, S., Farhadi, M., Biglari, H., Dastoorpour, M., Hashemzadeh, B., 2019. Ambient particulate matter concentration levels of Ahvaz, Iran. In: 2017. Environmental Geochemistry and Health, vol. 41, pp. 841–849.
Gupta, S., Raghuwanshi, G.S., Chanda, A., 2020. Effect of weather on COVID-19 spread in the US: a prediction model for India in 2020. Sci. Total Environ. 728, 138860.
Han, E., Tan, M.M.J., Turk, E., Sridhar, D., Leung, G.M., Shibuya, K., Asgari, N., Oh, J., García-Basteiro, A.L., Hanefeld, J., 2020. Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. Lancet.
Hao, X., Cheng, S., Wu, D., Wu, T., Lin, X., Wang, C., 2020. Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. Nature 584, 420–424.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780.
Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S., 1982. Rapid evolution of RNA genomes. Science 215, 1577–1585.
Imai, N., Cori, A., Dorigatti, I., Baguelin, M., Donnelly, C.A., Riley, S., Ferguson, N.M., 2020. Report 3: Transmissibility of 2019-nCoV. Imperial College London.
Jahangiri, M., Jahangiri, M., Najafgholipour, M., 2020. The sensitivity and specificity analyses of ambient temperature and population size on the transmission rate of the novel coronavirus (COVID-19) in different provinces of Iran. Sci. Total Environ. 728, 138872.
Jamshidi, S., Baniasad, M., Niyogi, D., 2020. Global to USA county scale Analysis of weather, urban density, mobility, homestay, and mask use on COVID-19. Int. J. Environ. Res. Publ. Health 17, 7847.
Jin, Q., Crippa, P., Pryor, S., 2020. Spatial characteristics and temporal evolution of the relationship between PM2. 5 and aerosol optical depth over the eastern USA during 2003–2017. Atmos. Environ. 239, 117718.
Koken, P.J., Piver, W.T., Ye, F., Elixhauser, A., Olsen, L.M., Portier, C.J., 2003. Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. Environ. Health Perspect. 111, 1312–1317.
Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner, N., Giorgi, E.E., Bhattacharya, T., Foley, B., 2020. Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182, 812–827 e819.
Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., du Plessis, L., Faria, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368, 493.
Krishna, R.K., Ghude, S.D., Kumar, R., Beig, G., Kulkarni, R., Nivdange, S., Chate, D., 2019. Surface PM2. 5 estimate using satellite-derived aerosol optical depth over India. Aerosol Air Q. Res. 19, 25–37.
Kumar, A., Nayar, K.R., Koya, S.F., 2020. COVID-19: challenges and its consequences for rural health care in India. Pub. Halth Practice 1.
Leung, K., Wu, J.T., Liu, D., Leung, G.M., 2020. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet 395, 1382–1393.
Lotfi, M., Hamblin, M.R., Rezaei, N., 2020. COVID-19: transmission, prevention, and potential therapeutic opportunities. Clin. Chim. Acta.
Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J., Luo, B., 2020. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. Sci. Total Environ. 724, 138226.
Mallires, K.R., Wang, D., Tipparaju, V.V., Tao, N., 2019. Developing a low-cost wearable personal exposure monitor for studying respiratory diseases using metal–oxide sensors. IEEE Sensor. J. 19, 8252–8261.
McCarty, J., Kaza, N., 2015. Urban form and air quality in the United States. Landsc. Urban Plann. 139, 168–179.
Mukandavire, Z., Nyabadza, F., Malunguza, N.J., Cuadros, D.F., Shiri, T., Musuka, G., 2020. Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios. PLoS one 15, e0236003.
Nishiyama, A., Wakasugi, N., Kirikae, T., Quy, T., Ha, L.D., Ban, V., 2008. Risk factors for SARS infection within hospitals in Hanoi, Vietnam. Jpn. J. Infect. Dis. 61, 388–390.
Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K.E., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N.F., Cattarino, L., Cooper, L.V., 2021. Reduction in mobility and COVID-19 transmission. Nat. Commun. 12, 1–9.
Pan, A., Liu, L., Wang, C., Guo, H., Hao, X., Wang, Q., Huang, J., He, N., Yu, H., Lin, X., 2020. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. Jama 323, 1915–1923.
Plante, J.A., Liu, Y., Liu, J., Xia, H., Johnson, B.A., Lokugamage, K.G., Zhang, X., Muruato, A.E., Zou, J., Fontes-Garfias, C.R., 2020. Spike mutation D614G alters SARS-CoV-2 fitness. Nature 1–6.
Rahman, M., Islam, M., Shimanto, M., Ferdous, J., Rahman, A.A.-N.S., Sagor, P., Esrat, T., 2020. Temperature Extreme May Exaggerate the Mortality Risk of COVID-19 in the Low- and Middle-Income Countries: A Global Analysis.
Rizwanul Fattah, I.M., Islam, A., Rahman, M.M., Chowdhury, M., 2020. Relationship between Climate Variables and New Daily COVID-19 Cases in Dhaka. Bangladesh).
Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., Sejdinovic, D., 2019. Detecting and quantifying causal associations in large nonlinear time series datasets. Sci. Adv. 5, eaau4996.
Sajadi, M.M., Habibzadeh, P., Vintzileos, A., Shokouhi, S., Miralles-Wilhelm, F., Amoroso, A., 2020. Temperature and Latitude Analysis to Predict Potential Spread and Seasonality for COVID-19. Available at: SSRN 3550308.
Salam, A., 2020. Impact and Correlation of Air Quality and Climate Variables with COVID-19 Morbidity and Mortality in Dhaka (Bangladesh. medRxiv).
Singh, M., Singh, M., Singh, B.B., Singh, R., Upendra, B., Kaur, R., Gill, S.S., Biswas, M.S., 2021. Quantifying COVID-19 enforced global changes in atmospheric pollutants using cloud computing based remote sensing. Remote Sensting App. Soc. Environ.: Soc. Environ. 22, 100489.
Stoecklin, S.B., Rolland, P., Silue, Y., Mailles, A., Campese, C., Simondon, A., Mechain, M., Meurice, L., Nguyen, M., Bassi, C., 2020. First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. Euro Surveill. 25, 2000094.
Tang, J.W., Liebner, T.J., Craven, B.A., Settles, G.S., 2009. A schlieren optical study of the human cough with and without wearing masks for aerosol infection control. J. R. Soc. Interface 6, S727–S736.
Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M.U.G., Li, B., Cai, J., Xu, B., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O.N., Yang, R., Grenfell, B.T., Pybus, O.G., Dye, C., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 368, 638.
Tosepu, R., Gunawan, J., Effendy, D.S., Lestari, H., Bahar, H., Asfian, P., 2020. Correlation between weather and covid-19 pandemic in Jakarta, Indonesia. Sci. Total Environ. 138436.
van Dorp, L., Richard, D., Tan, C.C., Shaw, L.P., Acman, M., Balloux, F., 2020. No evidence for increased transmissibility from recurrent mutations in SARS-CoV-2. Nat. Commun. 11, 1–8.
Volz, E., Hill, V., McCrone, J.T., Price, A., Jorgensen, D., O'Toole, Á., Southgate, J., Johnson, R., Jackson, B., Nascimento, F.F., 2020. Evaluating the effects of SARS-CoV-2 Spike mutation D614G on transmissibility and pathogenicity. Cell.
Wang, J., Pan, L., Tang, S., Ji, J.S., Shi, X., 2020. Mask use during COVID-19: a risk adjusted strategy. Environ. Pollut. 115099.
Wu, X., Nethery, R.C., Sabath, M., Braun, D., Dominici, F., 2020. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. Sci. Adv. 6, eabd4049.
Xie, J., Zhu, Y., 2020. Association between ambient temperature and COVID-19 infection in 122 cities from China. Sci. Total Environ. 724, 138201.
Yang, L., Tian, D., Liu, W., 2020. Strategies for Vaccine Development of COVID-19. Sheng Wu Gong Cheng Xue Bao, pp. 593–604.
Zhang, T., Zhao, G., Luo, L., Li, Y., Shi, W., 2020. Associations between Ambient Air Pollutants Exposure and Case Fatality Rate of COVID-19: a Multi-City Ecological Study in China.
Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273.
Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med.