Effects of Demographic and Weather Parameters on COVID-19 Basic Reproduction Number

Igor Salom[1], Andjela Rodic[2], Ognjen Milicevic[3], Dusan Zigic[1], Magdalena Djordjevic[1] and Marko Djordjevic[2]*

[1] Institute of Physics Belgrade, National Institute of the Republic of Serbia, University of Belgrade, Belgrade, Serbia,
[2] Quantitative Biology Group, Faculty of Biology, University of Belgrade, Belgrade, Serbia, [3] Department for Medical Statistics and Informatics, School of Medicine, University of Belgrade, Belgrade, Serbia

It is hard to overstate the importance of a timely prediction of the COVID-19 pandemic progression. Yet, this is not possible without a comprehensive understanding of environmental factors that may affect the infection transmissibility. Studies addressing parameters that may influence COVID-19 progression relied on either the total numbers of detected cases and similar proxies (which are highly sensitive to the testing capacity, levels of introduced social distancing measures, etc.), and/or a small number of analyzed factors, including analysis of regions that display a narrow range of these parameters. We here apply a novel approach, exploiting widespread growth regimes in COVID-19 detected case counts. By applying nonlinear dynamics methods to the exponential regime, we extract basic reproductive number R_0 (i.e., the measure of COVID-19 inherent biological transmissibility), applying to the completely naïve population in the absence of social distancing, for 118 different countries. We then use bioinformatics methods to systematically collect data on a large number of potentially interesting demographics and weather parameters for these countries (where data was available), and seek their correlations with the rate of COVID-19 spread. While some of the already reported or assumed tendencies (e.g., negative correlation of transmissibility with temperature and humidity, significant correlation with UV, generally positive correlation with pollution levels) are also confirmed by our analysis, we report a number of both novel results and those that help settle existing disputes: the absence of dependence on wind speed and air pressure, negative correlation with precipitation; significant positive correlation with society development level (human development index) irrespective of testing policies, and percent of the urban population, but absence of correlation with population density *per se*. We find a strong positive correlation of transmissibility on alcohol consumption, and the absence of correlation on refugee numbers, contrary to some widespread beliefs. Significant tendencies with health-related factors are reported, including a detailed analysis of the blood type group showing consistent tendencies on Rh factor, and a strong positive correlation of transmissibility with cholesterol levels. Detailed comparisons of obtained results with previous findings, and limitations of our approach, are also provided.

Keywords: COVID-19 transmissibility, environmental factors, basic reproduction number, COVID-19 demographic dependence, COVID-19 weather dependence

INTRODUCTION

The ancient wisdom teaches us that "knowing your adversary" is essential in every battle—and this equally applies to the current global struggle against the COVID-19 pandemic. Understanding the parameters that influence the course of the pandemic is of paramount importance in the ongoing worldwide attempts to minimize the devastating effects of the virus which, to the present moment, has already taken a toll of more than a million lives (Dong et al., 2020), and resulted in double-digit recession among some of the major world economies (World Bank, 2020a). Of all such factors, the ecological ones (both abiotic such as meteorological factors and biotic such as demographic and health-related population properties) likely play a prominent role in determining the dynamics of disease progression (Qu et al., 2020).

However, making good estimates of the effects that general demographic, health-related, and weather conditions, have on the spread of COVID-19 infection is beset by many difficulties. First of all, these dependencies are subtle and easily overshadowed by larger-scale effects. Furthermore, as the effective rate of disease spread is an interplay of numerous biological, medical, social, and physical factors, a particular challenge is to differentiate the dominating effects of local COVID-19-related policies, which are both highly heterogeneous and time-varying, often in an inconsistent manner. And this is precisely where, in our view, much of the previous research on this subject falls short.

There are not many directly observable variables that can be used to trace the progression of the epidemics on a global scale (i.e., for a large number of diverse countries). The most obvious one—the number of detected cases—is heavily influenced both by the excessiveness of the testing (which, in turn, depends on non-uniform medical guidelines, variable availability of testing kits, etc.) and by the introduced infection suppression measures (where the latter are not only non-homogeneous but are also erratically observed (Cohen and Kupferschmidt, 2020). Nevertheless, the majority of the research aimed to establish connections of the weather and/or demographic parameters with the spread of COVID-19 seeks correlations exactly with the raw number of detected cases (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Gupta et al., 2020; Iqbal et al., 2020; Li et al., 2020; Pourghasemi et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020). For the aforementioned reasons, the conclusions reached in this way are questionable. Other variables that can be directly measured, such as the number of hospitalized patients or the number of COVID-19 induced deaths (Pranata et al., 2020; Tosepu et al., 2020; Ward, 2020), again depend on many additional parameters that are difficult to take into account: level of medical care and current hospital capacity, advancements, and changing practices in treating COVID-19 patients, the prevalence of risk groups, and even on the diverging definitions of when hospitalization or death should be attributed to the COVID-19 infection. As such, these variables are certainly not suitable as proxies of the SARS-CoV-2 transmissibility *per se*.

On the other hand, as we here empirically find [and as theoretically expected (Anderson and May, 1992; Keeling and

Rohani, 2011)] the initial stage of the COVID-19 epidemic (in a given country or area) is marked by a period of a nearly perfect exponential growth for a wide range of countries, which typically lasts for about 2 weeks (based on our analysis of the available data). One can observe widespread dynamical growth patterns for many countries, with a sharp transition between exponential, superlinear (growth faster than linear), and sublinear (growth slower than linear) regimes (see **Figure 1**)—the last two representing a subexponential growth. We here concentrate on the initial exponential growth of the detected-case data (marked in red in **Figure 1**), characterizing the period *before* the control measures took effect, and with a negligible fraction of the population resistant to infection. Note that dates which correspond to the exponential growth regime (included in **Supplementary Table 1**) are different for each country, corresponding to the different start of COVID-19 epidemic in those countries.

We use the exponential growth regime to deduce the basic reproduction number R_0 (Martcheva, 2015), following a simple and robust mathematical (dynamical) model presented here. R_0 is a straightforward and important epidemiological parameter characterizing the inherent biological transmissibility of the virus, in a completely naïve population, and the absence of social distancing measures (Bar-On et al., 2020; Eubank et al., 2020). To emphasize the absence of social distancing in the definition (and inference) of R_0 used here, the term $R_{0,free}$ is also used, — for simplicity, we further denote $R_0 \equiv R_{0,free}$. R_0 is largely independent of the implemented COVID-19 policies and thus truly reflects the characteristics of the disease itself, as it starts to spread unhampered through the given (social and meteorological) settings. Namely, the exponential period ends precisely when the effect of control measures kick in, which happens with a delay of ~10 days after their introduction
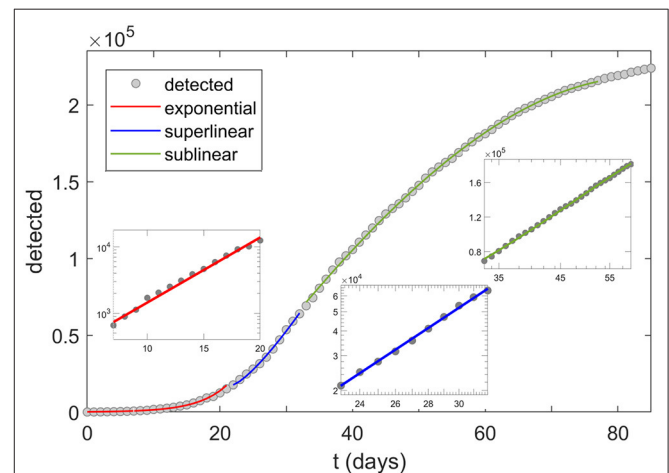
FIGURE 1 | COVID-19 growth regimes. Transitions of the growth patterns (here shown for Italy) from exponential (red), to superlinear (blue) and sublinear (green) regime. The three insets correspond to the log-linear scale (exponential), log-log scale (superlinear), and linear-log scale (sublinear). Dots correspond to detected infections, starting from 20.02.2020. In this study, $R_{0,free}$ is extracted from the slope of the first (exponential, i.e., log-linear) inset, corresponding to dates 29.02–13.03 in the case of Italy.

(The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020), corresponding to the disease latent period, and to the time between the symptom onset and the disease confirmation. Not only that very few governments had enacted any social measures before the occurrence of a substantial number of cases (Cohen and Kupferschmidt, 2020), but also the length of the incubation period makes it likely that the infection had been already circulating for some time through the community even before the first detected case (and that the effects of the measures are inescapably delayed in general). Also, the transition from the exponential to the subsequent subexponential phase of the epidemics is readily visible in the COVID data (see **Figure 1**). Furthermore, R_0 is invariant to the particular testing guidelines, as long as these do not significantly vary over the (here relatively short) studied period. Note that in **Figure 1** cumulative number of positive cases (also known as cumulative infection incidence) is shown, which has to monotonically increase—though with a decreasing rate, once the infection starts to slow down, i.e., once the subexponential growth (sublinear and superlinear regimes) is reached.

In the analysis presented here, we consider 42 different weather, demographic, and health-related population factors, whose analyzed ranges correspond to their variations exhibited in 118 world countries (not all of the parameters were available for all of the countries, as discussed in Section "Demographic and Weather Data Acquisition"). While some authors prefer more coherent data samples to avoid confusing effects of too many different factors (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020; Tosepu et al., 2020), this consideration is outweighed by the fact that large ranges of the analyzed parameters serve to amplify the effects we are seeking to recognize and to more reliably determine the underlying correlations. For example, while the value of the Human Development Index (HDI, a composite index of life expectancy, education, and per capita income indicators) varies from 0.36 to 0.96 over the set of analyzed countries, this range would drop by an order of magnitude (Global Data Lab, 2020) if the states of the US were chosen as the scope of the study (other demographic parameters exhibit similar behavior). The input parameters must take values in some substantial ranges to have measurable effects on R_0 (i.e., small variations may lead to effects that are easily lost in statistical fluctuations).

The number of considered parameters is also significant, especially when compared to other similar studies (Adhikari and Yin, 2020; Copat et al., 2020; Fareed et al., 2020; Iqbal et al., 2020; Rashed et al., 2020; Rychter et al., 2020; Singh and Agarwal, 2020; Thangriyal et al., 2020; Tosepu et al., 2020). In a model where a large number of factors are analyzed under the same framework, consistency of the obtained results, in terms of agreement with other studies, common-sense expectations, and their self-consistency, becomes an important check of applied methodology and analysis. Furthermore, a comprehensive and robust analysis is expected to generate new findings and lead to novel hypotheses on how environmental factors influence COVID-19 spread. Overall, we expect that the understanding achieved here will contribute to the ability to understand the behavior of the pandemics in the future and, by the same token,

to timely and properly take measures in an attempt to ameliorate the disease effects.

MODEL AND PARAMETER EXTRACTION

Modified SEIR Model and Relevant Approximations

There are various theoretical models and tools used to investigate and predict the progress of an epidemic (Keeling and Rohani, 2011; Martcheva, 2015). We here opted for the SEIR compartmental model, up to now used to predict or explain different features of COVID-19 infection dynamics (Maier and Brockmann, 2020; Maslov and Goldenfeld, 2020; Perkins and España, 2020; Tian et al., 2020; Weitz et al., 2020). The model is sufficiently simple to be applied to a wide range of countries while capturing all the features of COVID-19 progression relevant for extracting the R_0 values. The model assumes dividing the entire population into four (mutually exclusive) compartments with labels: (S)usceptible, (E)xposed, (I)fected, and (R)ecovered.

The dynamics of the model (which considers gradual transitions of the population from one compartment to the other) directly reflects the disease progression. Initially, a healthy individual has no developed SARS-CoV-2 virus immunity and is considered as "susceptible." Through contact with another infected individual, this person may become "exposed"— denoting that the transmission of the virus has occurred, but the newly infected person at this point has neither symptoms nor can yet transmit the disease. An exposed person becomes "infected"—in the sense of becoming contagious—on average after the so-called "latent" period which is, in the case of COVID-19, approximately 3 days. After a certain period of the disease, this person ceases to be contagious and is then considered as "recovered" (from the mathematical perspective of the model, "recovered" are all individuals who are no longer contagious, which therefore also includes deceased persons). In the present model, the recovered individuals are taken to be no longer susceptible to new infections (irrespectively of whether the COVID-19 immunity is permanent or not, it is certainly sufficiently long in the context of our analysis).

Accordingly, almost the entire population initially belongs to the susceptible class. Subsequently, parts of the population become exposed, then infected, and finally recovered. SARS-CoV-2 epidemic is characterized by a large proportion of asymptomatic cases (or cases with very mild symptoms) (Day, 2020), which leads to a large number of cases that remain undiagnosed. For this reason, only a portion of the infected will be identified (diagnosed) in the population, and we classify them as "detected." This number is important since it is the only direct observable in our model, i.e., the only number that can be directly related to the actual COVID-19 data.

This dynamic is schematically represented in **Figure 2**, and is governed by the following set of differential equations:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \tag{1}$$

$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E \tag{2}$$
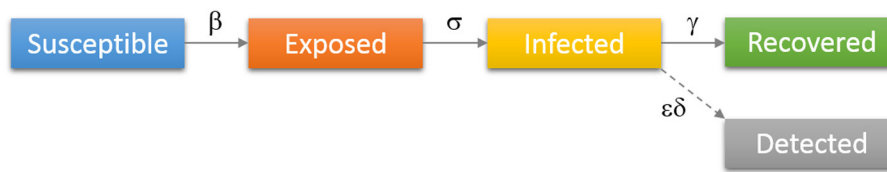
FIGURE 2 | Diagrammatic representation of the SEIR model with the added class of "Detected" patients. Individuals move (denoted by solid arrows) from Susceptible to Exposed to Infected to Recovered, with the rates indicated above arrows in the figure. Some of the infected are detected (diagnosed/confirmed), indicated by the dashed arrow.

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

$$\frac{dD}{dt} = \varepsilon \delta I \tag{5}$$

In the above equations, S, E, I, and R denote numbers of individuals belonging to, respectively, susceptible, exposed, infected, and recovered compartments, D is the cumulative number of detected cases, while N is the total population. Parameter β denotes the transmission rate, which is proportional to the probability of disease transmission in contact between a susceptible and an infectious subject. Incubation rate σ determines the rate at which exposed individuals become infected and corresponds to the inverse of the average incubation period. Recovery rate γ determines the transition rate between infected and recovered parts of the population, (i.e., $1/\gamma$ is the average period during which an individual is infectious). Finally, ε and δ are detection efficiency and the detection rate. All these rate parameters are considered constant during the analyzed (brief) period. Also, note that the constants in our model do not correspond to transition probabilities *per se*, but rather to transition rates (with units 1/time), so that e.g. γ and $\varepsilon\delta$ do not add to one. While rates in the model can be rescaled and normalized to directly correspond to transition probabilities, our formulation (with rates rather than probabilities) is rather common (see e.g., Keeling and Rohani, 2011), and also has a direct intuitive interpretation, where the transition rates correspond to the inverse of the period that individuals spend in a given compartment (see e.g., the explanation for γ above).

In the first stage of the epidemic, when essentially the entire population is susceptible (i.e., $S/N \approx 1$) and no distancing measures are enforced, the average number of secondary infections, caused directly by primary-infected individuals, corresponds to the basic reproduction number R_0. The infectious disease can spread through the population only when $R_0 > 1$ (Khajanchi et al., 2020a), and in these cases, the initial growth of the infected cases is exponential. Though R_0 is a characteristic of the pathogen, it also depends on environmental abiotic (e.g., local weather conditions), as well as biotic factors (e.g., prevalence of health conditions, and population mobility tightly related to the social development level).

Note that, as we seek to extract the basic reproduction number R_0 from the model for a wide range of countries,

the social distancing effects are not included in the model presented above. That is, the introduced model serves only to explain the exponential growth phase—note that this growth regime characterizes part of the infection progression where the social distancing interventions still did not take effect, and where the fraction of resistant (non-susceptible) population is still negligible. It is only this phase which is relevant for extracting R_0 that is used in the subsequent analysis. R_0 should not be confused with the effective reproduction number R_e, which takes into account also the effects of social distancing interventions and the decrease in the number of susceptibles due to acquired infection resistance. R_e is not considered in this work, as we are concerned with the factors that affect the inherent biological transmissibility of the virus, independently from the applied measures. That is, by considering R_0 rather than R_e, we disentangle the influence of meteorological and demographic factors on transmissibility (the goal of this study), from the effects of social distancing interventions (not analyzed here). The model can, however, be straightforwardly extended to include social distancing measures, as we did in (Djordjevic et al., 2020)—social distancing measures were also included through other frameworks (Khajanchi and Sarkar, 2020; Maier and Brockmann, 2020; Maslov and Goldenfeld, 2020; Perkins and España, 2020; Samui et al., 2020; Sarkar et al., 2020; Tian et al., 2020; Weitz et al., 2020). Such extensions are needed to explain the subexponential growth that emerges due to intervention measures (i.e., superlinear and sublinear growth regimes that are illustrated in **Figure 1** for Italy but are common for other countries as well).

COVID-19 Growth Regimes

If we observe the number of total COVID-19 cases (e.g., in a given country) as a function of time, there is a regular pattern that we observe: the growth of the detected COVID cases is initially exponential but slows down after some time—when we say it enters the subexponential regime. The subexponential regime can be further divided into the superlinear (growing asymptotically faster than a linear function) and sublinear regime (the growth is asymptotically slower than a linear function). This typical behavior is illustrated, in the case of Italy, in **Figure 1** above. The transition to the subexponential regime occurs relatively soon, much before a significant portion of the population gains immunity, and is a consequence of the introduction of the infection suppression measures.

Inference of the Basic Reproduction Number R_0

In the initial exponential regime, a linear approximation to the model can be applied. Namely, in this stage, almost the entire population is susceptible to the virus, i.e., S/N≈1, which simplifies the Equation (2) to:

$$\frac{dE}{dt} = -\sigma E + \beta I. \tag{6}$$

By combining expressions (3) and (6) one obtains:

$$\frac{d}{dt}\begin{pmatrix} E \\ I \end{pmatrix} = \begin{pmatrix} -\sigma & \beta \\ \sigma & -\gamma \end{pmatrix} \begin{pmatrix} E \\ I \end{pmatrix} = A \begin{pmatrix} E \\ I \end{pmatrix}, \tag{7}$$

where we have introduced a two-by-two matrix:

$$A = \begin{pmatrix} -\sigma & \beta \\ \sigma & -\gamma \end{pmatrix} \tag{8}$$

The solution for the number of infected individuals can now be written:

$$I(t) = C_1 \cdot e^{\lambda_+ t} + C_2 \cdot e^{\lambda_- t}, \tag{9}$$

where λ_+ and λ_- denote eigenvalues of the matrix A, i.e., the solutions of the equation:

$$\det(A - \lambda I) = 0. \tag{10}$$

The eigenvalues must satisfy:

$$\begin{vmatrix} -\sigma - \lambda & \beta \\ \sigma & -\gamma - \lambda \end{vmatrix} = 0,$$

leading to:

$$(\lambda + \sigma) \cdot (\lambda + \gamma) - \beta \cdot \sigma = 0. \tag{11}$$

The solutions of (11) are:

$$\lambda_\pm = \frac{-(\gamma + \sigma) \pm \sqrt{(\gamma - \sigma)^2 + 4\beta\sigma}}{2}. \tag{12}$$

Since $\lambda_- < 0$, the second term in (9) can be neglected for sufficiently large t. More precisely, numerical analysis shows that this approximation is valid already after the second day, while, for the extraction of R_0 value we will anyhow ignore all data before the fifth day (for the analyzed countries, numbers of cases before the fifth day were generally too low, hence this early data is dominated by stochastic effects/fluctuations). Hence, $I(t)$ is proportional to $exp(\lambda_+ t)$, i.e.:

$$I(t) = I(0) \cdot e^{\lambda_+ t}. \tag{13}$$

By using β from (12) and $R_0 = \frac{\beta}{\gamma}$ (Keeling and Rohani, 2011; Martcheva, 2015), we obtain:

$$R_0 = 1 + \frac{\lambda_+ \cdot (\gamma + \sigma) + \lambda_+^2}{\gamma \cdot \sigma}. \tag{14}$$

From (13) and (5) we compute:

$$D(t) = \varepsilon \cdot \delta \cdot I(0) \cdot \frac{(e^{\lambda_+ t} - 1)}{\lambda_+}. \tag{15}$$

By taking the logarithm, the above expression leads to:

$$\log(D(t)) = \log(\varepsilon \delta I(0)/\lambda_+) + \lambda_+ \cdot t, \tag{16}$$

from which λ_+ can be obtained as the slope of the $\log(D(t))$ function. From Equation (14), we thus obtain the R_0 value as a function of the slope of $\log(D(t))$, where the latter can be efficiently inferred from the plot of the number of detected COVID-19 cases for a large set of countries.

The SEIR model and the above derivation of R_0 assume that the population belonging to different compartments is uniformly mixed. Possible heterogeneities may tend to increase R_0 values (Keeling and Rohani, 2011). However, this would not influence the results obtained below, as our R_0 values are consistently inferred for all analyzed countries by using the same model, methodology, and parameter set. Moreover, our R_0 values are in agreement with the prevailing estimates in the literature (Najafimehr et al., 2020).

Demographic and Weather Data Acquisition

For the countries for which R_0 was determined through the procedure above, we also collect a broad spectrum of meteorological and demographic parameter values. Overall, 118 countries were selected for our analysis, based on the relevance of the COVID-19 epidemiological data. Namely, a country was considered as relevant for the analysis if the number of detected cases on June 15th was higher than a threshold value of 1,000. A few countries were then discarded from this initial set, where the case count growth was too irregular to extract any results, possibly due to inconsistent or irregular testing policies. As a source for detected cases, we used (World Bank, 2020b; Worldometer, 2020).

In the search for factors correlated with COVID-19 transmissibility, we have analyzed overall 42 parameters, 11 of which are related to weather conditions, 30 to demographics or health-related population characteristics, and one parameter quantifying a delay in the epidemic's onset (data provided in **Supplementary Tables 2–5**). Not all of these parameters were available for all of the considered countries. In particular, data on the prevalence of blood types (**Supplementary Table 4** in the Supplement) was possible to find for 83 of the 118 countries, while, primarily due to scarce data on pollutant concentrations during the epidemics, almost 30% of entries in **Supplementary Table 5** in Supplement had to be left blank for this category. Nevertheless, we opted to include these parameters in our report: despite the lower number of values, some of these parameters exhibited strong and highly statistically significant correlations with R_0, warranting their inclusion.

Our main source of weather data was project POWER (Prediction of Worldwide Energy Resources) of the NASA agency (NASA Langley Research Center, 2020). A dedicated

Python script was written and used to acquire weather data via the provided API (Application Programming Interface). NASA project API allows a large set of weather parameters to be obtained for any given location (specified by latitude and longitude) and given date (these data are provided in the **Supplementary Table 7**). From this source, we gathered data on temperature (estimated at 2 m above ground), specific humidity (estimated at 2 m above ground), wind speed (estimated at 2 m above ground), and precipitation (defined as the total column of precipitable water). Data on air pressure (at ground level) and UV index (international standard measurement of the strength of sunburn-producing ultraviolet radiation) were collected via similar API from World Weather Online source (World Weather Online, 2020), using the same averaging methodology. Since we needed to assign a single value to each country (for each analyzed parameter), the following method was used for averaging meteorological data. In each country, a number of largest cities[1] were selected and weather data was taken for the corresponding locations. These data was then averaged, weighted by the population of each city, followed by averaging over the period used for R_0 estimation (more precisely, to account for the time between disease transmission and the case confirmation, we shifted this period 12 days into the past). The applied averaging method used here can be of limited adequacy in countries spreading over multiple climate zones, but is still expected to provide reasonable single-value estimates of the weather parameters, particularly since the averaging procedure was formulated to reflect the most likely COVID-19 hotspots in a given country.

Demographic data was collected from several sources. Percentage of the urban population, refugees, net migration, social and medical insurance coverage, infant mortality, and disease (CVD, cancer, diabetes, and CRD) risk was taken from the World Bank organization (World Bank, 2020b). The HDI was taken from the Our World in Data source (Our World in Data, 2020), while median age information was obtained from the CIA website. The source of most of the considered medical parameters: cholesterol, raised blood pressure, obesity, inactivity, BSG vaccination as well as data on alcohol consumption and smoking prevalence was World Health Organization (World Health Organization, 2020). Data for blood types were taken from the Wikidata web site. BUCAP parameter, representing population density in the built-up area, was taken from GHS Urban Center Database 2015 (European Commission Global Human Settlement, 2020). The onset parameter, determining the delay (in days) of the epidemic's start, was inferred from COVID-19 counts data. We used the most recent available data for all the parameters.

RESULTS

The $\log(D(t))$ function, for a subset of selected countries, is shown in **Figure 3**. The obvious linear dependence confirms that

---

[1]This number was determined for each country by the following condition: the total population of the cities taken into consideration had to surpass 10 percent of the overall population of the country.

the progression of the epidemic in this stage is almost perfectly exponential. Note that our model exactly reproduces this early exponential growth (see Equation 13), happening under the assumption of a small fraction of the population being resistant, and the absence of the effect of social distancing interventions. From **Figure 3**, we see that this behavior, predicted by the model for the early stage of the epidemic, is also directly supported by the data, i.e., the exponential growth in the cumulative number of confirmed cases is indeed observed for a wide range of countries. For each country, the parameter λ_+ is directly obtained as the slope of the corresponding linear fit of the $\log(D(t))$, and the basic reproduction number R_0 is then calculated from Equation (14). Here, we used the following values for the incubation rate, $\sigma = 1/3$ day^{-1}, and for the recovery rate $\gamma = 1/4$ day^{-1}, per the commonly accepted values in the literature (Bar-On et al., 2020). Note that possible variations in these two values would not significantly affect any conclusions about R_0 correlations, due to the mathematical properties of the relation (14): it is a strictly monotonous function of λ_+ and the linear term $\lambda_+ \cdot (\gamma + \sigma)/\gamma \cdot \sigma$ dominantly determines the value of R_0.

Supplementary tables contain the values for 42 variables, for all countries. Correlations of each of the variables with R_0 are given in **Supplementary Table 6**. Values for the Pearson correlation coefficient are further shown below, though consistent conclusions are also obtained by Kendall and Spearman correlation coefficients (which do not assume a linear relationship between variables). Correlation coefficients were calculated in the usual manner: as the correlation of the vector of parameter values with the vector of R_0 values, by taking into account all available data (for parameters that were available across all of the countries, both of the vectors were 118 dimensional; if values were missing for certain countries, these countries were simply ignored and lower-dimensional vectors were compared).

The first set of results that corresponds to, roughly speaking, general demographic data, is presented in **Figure 4**. The plot in panel A shows the distribution of R_0 vs. HDI values for all countries, where a higher HDI score indicates the more prosperous country concerning life expectancy, education, and per capita income (Sagar and Najam, 1998). This parameter was included in the study due to a reasonable expectation that a higher level of social development also implies a higher level of population interconnectedness and mixing (stronger business and social activity, more travelers, more frequent contacts, etc.), and hence that HDI could be related to the SARS-CoV-2 transmissibility. Indeed, we note a strong, statistically highly significant correlation between the HDI and the R_0 value, with $R = 0.37$, and $p = 4 \cdot 10^{-5}$, demonstrating that the initial expansion of COVID-19 was faster in more developed societies.

The social security and health insurance coverage (INS) "shows the percentage of population participating in programs that provide old age contributory pensions (including survivors and disability) and social security and health insurance benefits (including occupational injury benefits, paid sick leave, maternity, and other social insurance)" (World Bank, 2020b). Reflecting the percentage of the population covered by medical insurance and likely feeling more protected from the financial
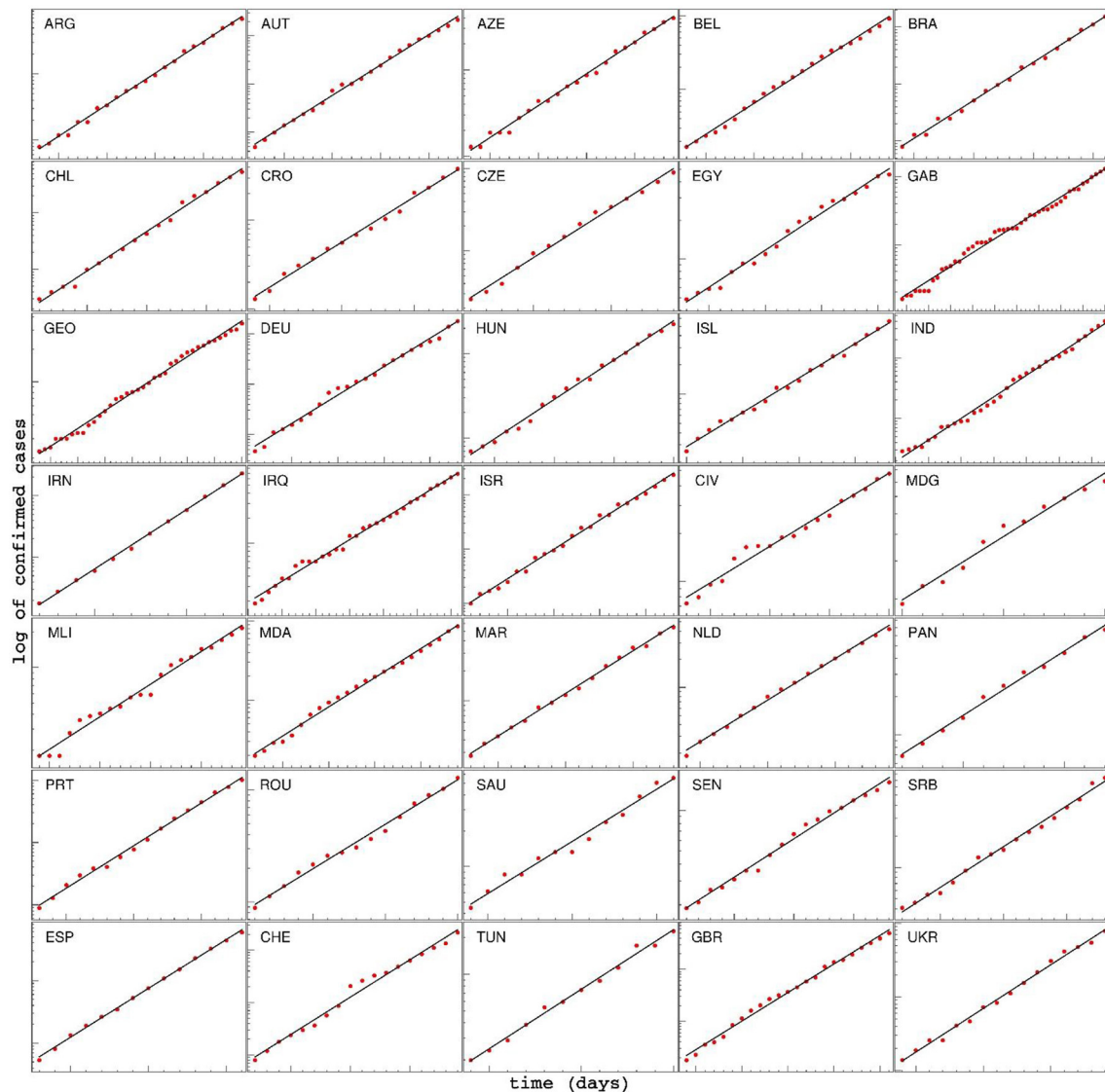
FIGURE 3 | Time dependence of the detected cases for various countries, during the initial period of the epidemic, shown on a log-linear scale. The linear fit of log(D) shows that the spread of COVID-19 in this phase is very well approximated by exponential growth. Note that the values on axes are chosen differently for each country, in order to emphasize the exponential character of the growth. For each country, the start and end dates of the exponential regime, together with the extracted slope λ_+, are provided in the **Supplementary Table 1**. ARG, Argentine; AUT, Austria; AZE, Azerbaijan; BEL, Belgium; BRA, Brazil; CHL, Chile; CRO, Croatia; CZE, Czech Republic; EGY, Egypt; GAB, Gabon; GEO, Georgia; DEU, Germany; HUN, Hungary; ISL, Iceland; IND, India; IRN, Iran; IRQ, Iraq; ISR, Israel; CIV, Cote d'Ivoire; MDG, Madagascar; MLI, Mali; MDA, Moldova; MAR, Morocco; NLD, Netherlands; PAN, Panama; PRT, Portugal; ROU, Romania; SAU, Saudi Arabia; SEN, Senegal; SRB, Serbia; ESP, Spain; CHE, Switzerland; TUN, Tunis; GBR, Great Britain; UKR, Ukraine.

effects of the epidemics, this indicator shows a strong ($R = 0.4$) and highly significant ($p = 4 \cdot 10^{-4}$) positive correlation with R_0. The percentage of urban population (UP) and BUCAP density (BAP) are both included as measures of how concentrated is the population of the country. While the UP value simply shows what percentage of the population lives in cities, the BUCAP parameter denotes the amount of built-up area per person. Of the two, the former shows a highly significant positive correlation with the COVID-19 basic reproduction number, whereas the latter shows no correlation. Median age (MA) should be of obvious potential relevance in COVID-19 studies since it is well

known that the disease more severely affects the older population (Jordan et al., 2020). Thus, we wanted to investigate also if there is any connection of age with the virus transmissibility. Our results are suggestive of such a connection, since we obtained a strong positive correlation of age with R_0, with very high statistical confidence. Infant mortality (IM) is defined as the number of infants dying before reaching 1 year of age, per 1,000 live births. Lower IM rates can serve as another indicator of the prosperity of a society, and it turns out that this measure is also strongly correlated, but negatively, with $R = -0.36$ and $p = 8 \cdot 10^{-5}$ (showing again that more developed countries, i.e.,

FIGURE 4 | (A) R_0 vs. HDI as an example of transmissibility dependence on demographic data (Pearson correlation denoted as R). **(B)** Pearson correlations of R_0 with (from left to right): social security and health insurance coverage (INS), percentage of urban population (UP), BUCAP measure of population density (BAP), median age (MA), infant mortality (IM), net migration (I-E), and percentage of refugee population by country or territory of asylum (RE). The statistical significance of each correlation is indicated in the legend, while "ns" stands for "no significance."
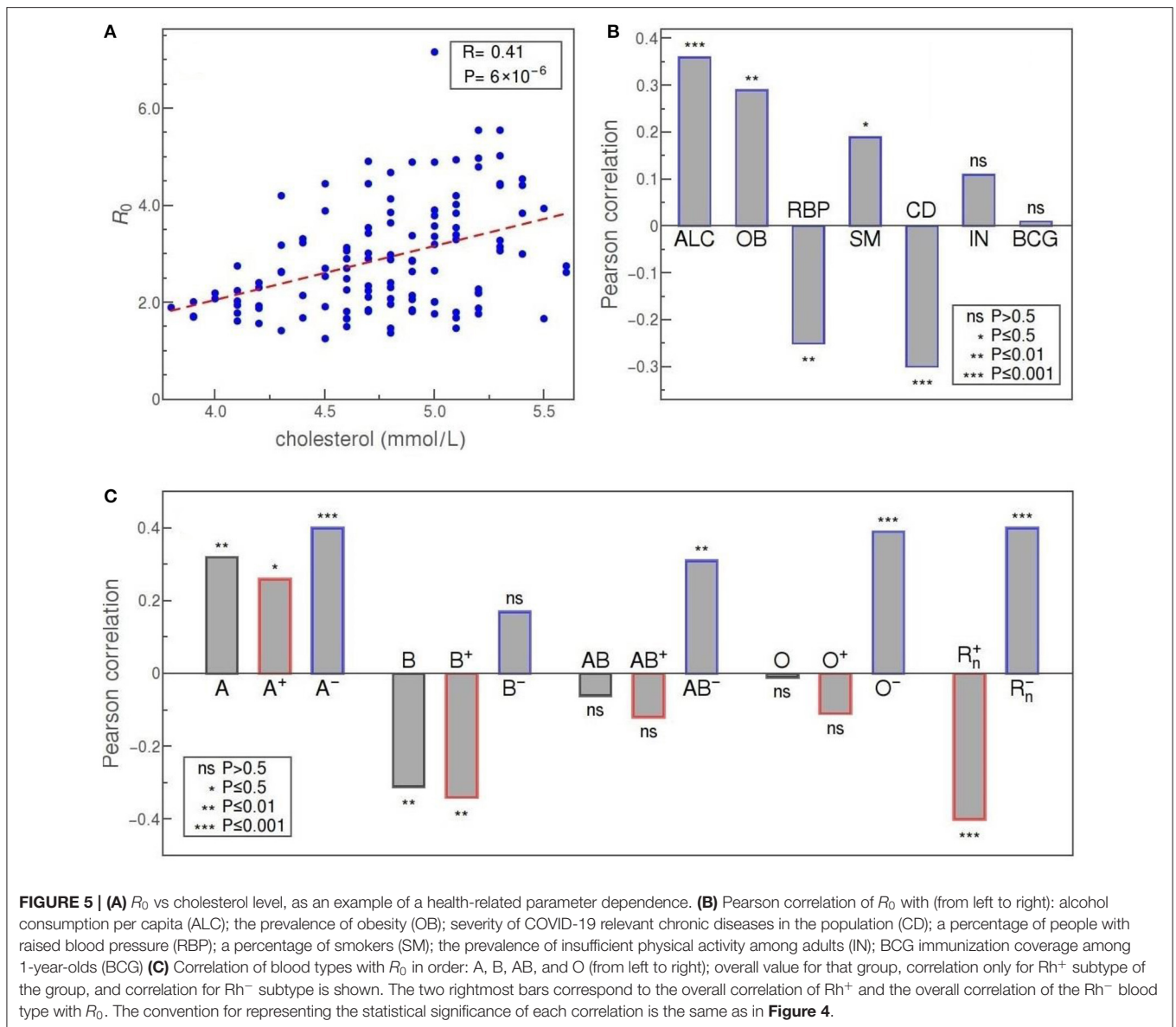
those with lower IM rates, have experienced more rapid spread of the virus infection). Net migration (I-E) represents the 5-year estimates of the total number of immigrants less the annual number of emigrants, including both citizens and non-citizens. This number, related to the net influx of foreigners, turns out to be positively correlated, in a statistically significant way, with R_0. However, according to our data, the percentage of refugees, defined as the percentage of the people in the country who are legally recognized as refugees and were granted asylum in that country, is not correlated with R_0 at all.

Another set of parameters corresponds to medically-related demographic parameters and is shown in the upper part of **Figure 5**. The plot in panel A represents the average blood cholesterol level (in mmol/L) in the population of various countries, plotted against the value of R_0. The two parameters are strongly correlated, with $R = 0.4$, and $p = 6 \cdot 10^{-6}$. Another demographic parameter with clear medical relevance, that has a comparatively strong and significant positive correlation with R_0, is the alcohol consumption per capita (ALC), as shown in panel B of **Figure 5**. Our data shows that R_0 is also positively correlated, with high statistical significance, with the prevalence of obesity and to a somewhat smaller extent with the percentage of smokers. Here, obesity is defined as having a body-mass index over 30. A medical parameter that is strongly, but negatively, correlated with R_0, is a measure of prevalence and severity of COVID-19 relevant chronic diseases in the population (CD). This parameter is defined as "the percent of 30-year-old-people who would die before their 70th birthday from any of cardiovascular disease, cancer, diabetes, or chronic respiratory disease, assuming that s/he would experience current mortality rates at every age and s/he would not die from any other cause of death" (World Bank, 2020b). The percentage of people with raised blood pressure (RBP) is also negatively

correlated with R_0, though this correlation is not as strong and as statistically significant as in the case of the CD parameter. Here, raised blood pressure is defined as systolic blood pressure over 140 or diastolic blood pressure over 90, in the population older than 18. The percentage of smokers exhibits statistically significant (though not large) positive correlation. Two medical-demographic parameters that show no correlation with R_0 in our data are the prevalence of insufficient physical activity among adults aged over 18 (IN) and BCG immunization coverage among 1-year-olds (BCG).

In **Figure 5C** we see that blood types are, in general, strongly correlated with R_0. The highest positive correlation is exhibited by A$^-$ and O$^-$ types, with a Pearson correlation of 0.4 and 0.39, and a very high statistical significance of $p = 10^{-4}$ and $p = 2 \cdot 10^{-4}$, respectively. Taken as a whole, group A is still strongly and positively correlated with R_0, albeit with a bit lower statistical significance (A$^+$ type correlation has p-value two orders of magnitude higher than A$^-$). This is not so for group O that, overall, does not seem to be correlated to R_0 (O$^+$ even shows a certain negative correlation but without statistical significance). Our data reveals a highly significant positive correlation also for AB$^-$ subtype ($R = 0.31$, $p = 0.003$), while neither the AB$^+$ subtype nor overall AB group is significantly correlated with the basic reproduction number. Clear negative correlation is exhibited only by B blood group ($R = -0.31$, $p = 0.004$), mostly due to the negative correlation of its B$^+$ subtype ($R = -0.34$, $p = 0.001$), whereas B$^-$ subtype is not significantly correlated with R_0 in our data. If we consider the rhesus factor alone, we again observe very strong correlations with R_0 and with very high statistical significance: Rh$^-$ and Rh$^+$ correlate positively ($R = 0.4$) and negatively ($R = -0.4$), respectively, with very high statistical significance ($p = 2 \cdot 10^{-4}$). The tendency of Rh$^-$ and Rh$^+$ to, respectively, increase and decrease the transmissibility,

FIGURE 5 | (A) R_0 vs cholesterol level, as an example of a health-related parameter dependence. **(B)** Pearson correlation of R_0 with (from left to right): alcohol consumption per capita (ALC); the prevalence of obesity (OB); severity of COVID-19 relevant chronic diseases in the population (CD); a percentage of people with raised blood pressure (RBP); a percentage of smokers (SM); the prevalence of insufficient physical activity among adults (IN); BCG immunization coverage among 1-year-olds (BCG) **(C)** Correlation of blood types with R_0 in order: A, B, AB, and O (from left to right); overall value for that group, correlation only for Rh$^+$ subtype of the group, and correlation for Rh$^-$ subtype is shown. The two rightmost bars correspond to the overall correlation of Rh$^+$ and the overall correlation of the Rh$^-$ blood type with R_0. The convention for representing the statistical significance of each correlation is the same as in **Figure 4**.

is therefore consistent with the results obtained for all four individual blood-groups.

In **Figure 6**, the onset represents the delay of the exponential phase and is defined, for each country, as the number of days from February 15 to the start of the exponential growth of detected cases. The motivation was to check for a possible correlation between the delay in the onset of the epidemic and the rate at which it spreads. Indeed, our data shows that such correlation exists and that it is strong and statistically significant: $R = -0.48$ and $p = 4 \cdot 10^{-8}$. In other words, the later the epidemic started, the lower (on average) is the basic reproduction number.
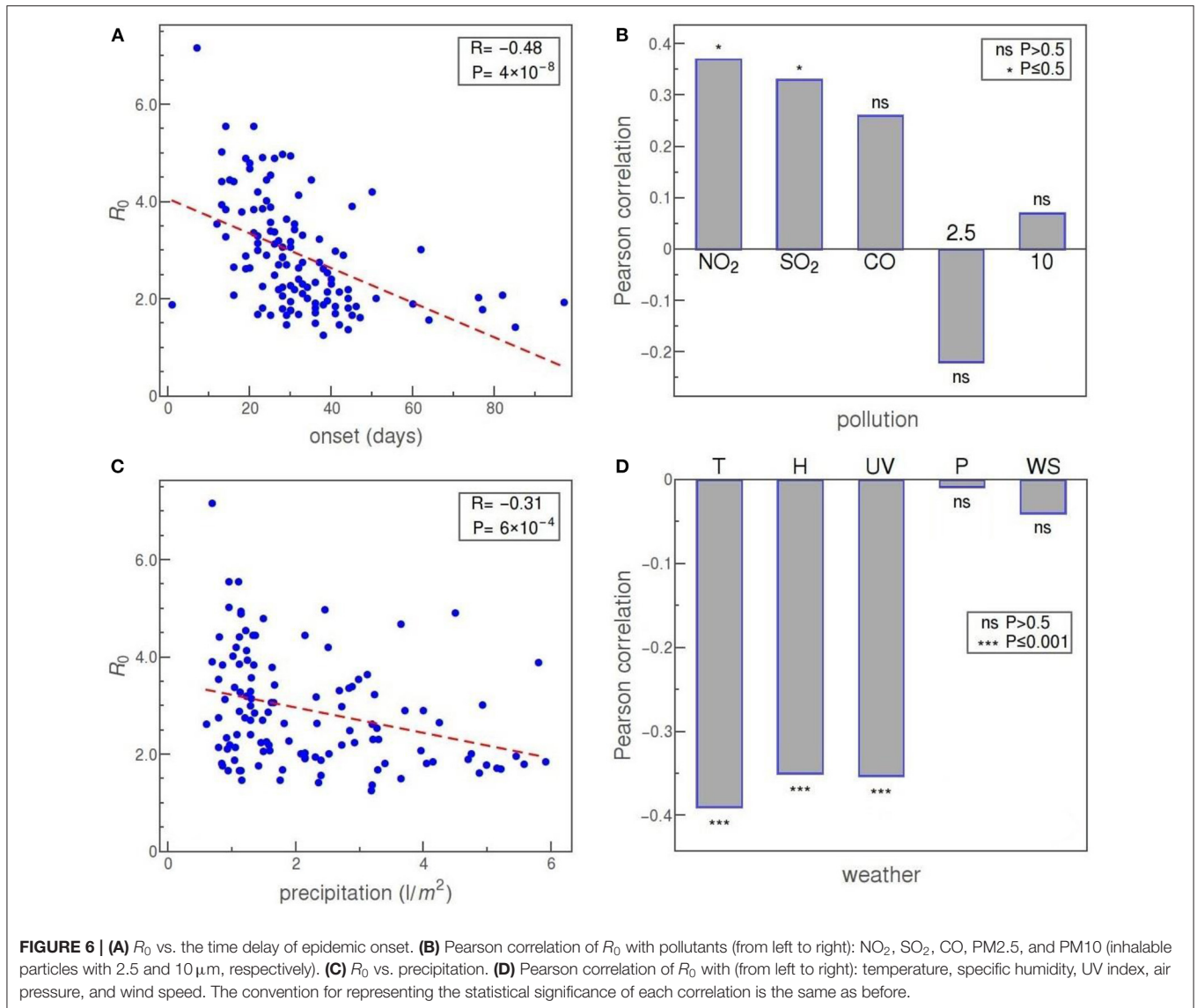
Panel B of **Figure 6** shows the correlation of R_0 with some of the commonly considered air pollutants. Our data reveal a statistically significant positive correlation of R_0 with NO_2 and SO_2 concentrations. Other pollutants—CO, PM2.5 (fine inhalable particles, with diameters that are generally 2.5 μm

and smaller), and PM10 (inhalable particles, up to 10 nm in diameter)—show no statistically significant correlation with R_0.

Next, we consider weather factors. Panels C and D of **Figure 6** show correlations of precipitation, temperature, specific humidity, UV index, air pressure, and wind speed with the reproduction number R_0. Of these, precipitation, temperature, specific humidity, and UV index show a strong negative correlation, at a high level of statistical significance. Of the other two parameters, both air pressure and wind speed are not correlated at all with R_0 in our data.

DISCUSSION

The present paper aimed to establish relations between the COVID-19 transmissibility and a large number of

FIGURE 6 | **(A)** R_0 vs. the time delay of epidemic onset. **(B)** Pearson correlation of R_0 with pollutants (from left to right): NO$_2$, SO$_2$, CO, PM2.5, and PM10 (inhalable particles with 2.5 and 10 µm, respectively). **(C)** R_0 vs. precipitation. **(D)** Pearson correlation of R_0 with (from left to right): temperature, specific humidity, UV index, air pressure, and wind speed. The convention for representing the statistical significance of each correlation is the same as before.

demographic and weather parameters. As a measure of COVID-19 transmissibility, we have chosen the basic reproduction number R_0—a quantity that is essentially independent of the variations in both the testing policies and the introduced social measures (as discussed in the Introduction), in distinction to many studies on transmissibility that relied on the total number of detected case counts [see e.g., (Adhikari and Yin, 2020; Correa-Araneda et al., 2020; Fareed et al., 2020; Gupta et al., 2020; Iqbal et al., 2020; Li et al., 2020; Pourghasemi et al., 2020; Rashed et al., 2020; Singh and Agarwal, 2020)]. We have covered a substantial number of demographic and weather parameters, and included in our analysis all world countries that were significantly affected by the COVID-19 pandemic (and had a reasonable consistency in tracking the early phase of infection progression). While a number of manuscripts have been devoted to factors that may influence COVID-19 progression, only a few used an estimate of R_0 or some of its

proxies (Coccia, 2020; Contini and Costabile, 2020; Copiello and Grillenzoni, 2020)—these studies were however limited to China, and included a small set of meteorological variables, with conflicting results obtained for their influence on R_0. Therefore, a combination of (i) using a reliable and robust measure of COVID-19 transmissibility, and (ii) considering a large number of factors that may influence this transmissibility within the same study/framework, distinguishes our study over prior work. We, however, must be cautious when it comes to further interpretation of the obtained data. As always, we must keep in mind that "correlation does not imply causation" and that further research is necessary to identify possible confounding factors and establish which of these parameters truly affect the COVID-19 transmissibility. Due to the sheer number of studied variables, an even larger number of parameters that might be relevant but are inaccessible to study (or even impossible to quantify), as well as due to possible intricate mutual relations of

the factors that may influence COVID-19 transmission, this is a highly non-trivial task. While we postpone any further analysis in this direction to future studies, we will, nevertheless, consider here the possible interpretations of the obtained correlations, assuming that they also probably indicate the existence of at least some causation. We below provide a detailed comparison between our results and previous findings. While a detailed discussion is presented, despite our best effort, we may have missed some of the relevant references due to an extremely rapidly developing field. Nevertheless, we point out a clear distinction of our work with previous studies, as outlined in this paragraph.

We will first consider the demographic variables presented in **Figure 4**. The obtained correlation of the HDI with the basic reproduction number is both strong and hardly surprising. The level of prosperity and overall development of a society is necessarily tied with the degree of population mobility and mixing, traffic intensity (in particular air traffic), business and social activity, higher local concentrations of people, and other factors that directly or indirectly increase the frequency and range of personal contacts (Gangemi et al., 2020), rendering the entire society more vulnerable to the spread of viruses. In this light, it is reasonably safe to assume that the obtained strong and highly statistically significant correlation of HDI with R_0 reflects a truly causal connection. However, some authors offer also a different explanation: that higher virus transmission in more developed countries is a consequence of more efficient detection of COVID-19 cases due to the better-organized health system (Gangemi et al., 2020)—but since our R_0 measure does not depend on detection efficiency, presented results can be taken as evidence against such hypothesis.

The interpretation is less clear for other demographic parameters, for example, the percentage of the population covered by medical and social insurance programs (INS). While there seem to be no previous studies discussing this parameter, one possibility is to attribute its strong positive correlation with R_0 to a hypothetical tendency of population to more easily indulge in the epidemiologically-risky behavior if they feel well-protected, both medically and financially, from the risks posed by the virus; conversely, that the population that cannot rely on professional medical care in the case of illness is likely to be more cautious not to contract the virus. The other is, of course, to see this correlation as an indirect consequence of the strong correlation of this parameter with HDI—which is also, almost certainly, the underlying explanation of the infant mortality (IM) correlation, where low mortality ratios point to a better medical system, which goes hand in hand with the overall prosperity and development of the country (Ruiz et al., 2015) (thus the negative correlation with R_0).

Similarly, the strong positive correlation of median age (MA) with R_0 might be a mere consequence of its clear relation with the overall level of development of the country (Gangemi et al., 2020), but it can be also considered in the light of the fact that clinical and epidemiological studies have unanimously shown that the elderly are at higher risk of developing a more severe clinical picture, and our result may indicate that the virus also spreads more efficiently in the elderly population. Possible

explanations may include: drugs frequently prescribed to this population that increase levels of ACE2 receptors (Shahid et al., 2020), a general weakening of the immune system with age leading to a greater susceptibility to viral infections (Pawelec and Larbi, 2008), and a large number of elderly people grouped in nursing homes, where the virus can expand very quickly (Kimball et al., 2020).

The correlation of population density with R_0, or the lack of thereof, is more challenging to explain. Naively, one could expect that COVID-19 spreads much more rapidly in areas with a large concentration of people, but, if exists, this effect is not that easily numerically captured. As the standard population density did not show any correlation with the reproduction number R_0 (not shown), we explored some more subtle variants. Namely, the simplest reason why the data shows no correlation of R_0 with population density would be that the density, calculated in the usual way, is too averaged out: the most densely-populated country on our list, Monaco, has roughly 10,000 times more people per square kilometer than the least densely-populated Australia. However, Melbourne downtown has a similar population density as Monaco and far more people, so one would expect no a priori reason that its infection progression would be slower (and the R_0 rate for Australia as a whole will be dominantly determined by the fastest exponential expansion occurring anywhere on its territory). For this reason, we included the BUCAP parameter into the analysis, which takes into account only population density in built-up areas. Surprisingly, even this parameter did not exhibit any statistically significant correlation. Actually, several studies may serve as examples showing that the correlation of population density with the rate of COVID-19 expansion can be expected only under certain conditions since the frequency of contacts between people is to a large extent modulated by additional geographical, economic, and sociological factors (Berg et al., 2020; Carozzi, 2020; Pourghasemi et al., 2020; Rashed et al., 2020). Our observed absence of a correlation could be therefore expected and possibly indicates that such a correlation should be sought at the level of smaller populated areas—for example, individual cities (Yu et al., 2020). This conclusion is somewhat supported by the obtained highly significant and strong positive correlation of R_0 with the percentage of the population living in cities (UP) and which probably reflects the higher number of encounters between people in a more densely populated, urban environment (Li et al., 2020). It is also possible that virus spread might have a highly non-linear dependence on the population density— namely, that an outbreak in a susceptible population requires a certain threshold value of its density, while below that value population density ceases to be a significant factor influencing virus (Scheffer, 2009; Carozzi, 2020; Coro, 2020).

Another demographic parameter that exhibits a significant correlation with R_0 in our data is the net migration (I-E), denoting the number of immigrants less the number of emigrants. Unlike this number, which shows a positive correlation, the number of refugees (RE) seems not to be correlated at all. By definition, migrants deliberately choose to move to improve their prospects, while refugees have to move to save their lives or preserve their freedom. Migrants (e.g., in

economic or academic migration), arguably tend to stay in closer contact with the country of their origin and have more financial means for that, which likely contributes to more frequent border crossings and more intensive passenger traffic (Fan et al., 2020), thereby promoting the infection spread. On the other hand, refugees are mostly stationed in refugee camps, there is less possibility of spreading the virus outside through contacts with residents, but there is a high possibility of escalation of the epidemic within camps with a high concentration of people (Hargreaves et al., 2020). We did not find any other attempt in the literature to examine this issue. In any case, our results demonstrate that refugees are certainly not a primary cause of concern in the pandemics, contrary to fears expressed in some media.

Of the medical factors, the strongest correlation of R_0 is established with elevated cholesterol levels, as shown in **Figure 5**. Cholesterol may be associated with a viral infection and further disease development through a complex network of direct and indirect effects. *In vitro* studies of the role of cholesterol in virus penetration into the host body, done on several coronaviruses, indicate that its presence in the lipid rafts of the cell membrane is essential for the interaction of the virus with the ACE2 receptor, and also for the latter endocytosis of the virus (Radenkovic et al., 2020). Obesity prevalence (OB) also exhibits a highly significant, though somewhat weaker correlation with R_0, which might be a consequence of the common connection between obesity and cholesterol: in principle, obesity might be a relevant factor in the COVID-19 epidemic exactly due to the effects of cholesterol on SARS-CoV-2 susceptibility. Of course, other effects might be at play, e.g., the fact that the adipose tissue of obese people excessively produces pro-inflammatory cytokines (Sattar et al., 2020). In the case of obesity, a simple explanation via relation to HDI is not available, since obesity does not show a simple correlation with the society development (Haidar and Cosman, 2011). Overall, while the correlation of obesity with a more severe prognosis in COVID-19 is well established in the literature, its relation to COVID-19 transmissibility is only mentioned in Li et al. (2020) and hitherto unexplained.

Often related to obesity is also raised blood pressure (RBP), and we have discovered that this factor is also correlated, at high statistical significance, with R_0. While this seems to be the first study correlating high blood pressure with the SARS-CoV-2 transmission rate, it is known that, based on clinical studies, RBP appears to be a risk factor for hospitalization and death due to COVID-19 (Ran et al., 2020a; Schiffrin et al., 2020). In this light, it might be surprising that the correlation between RBP and R_0 turns out to be negative. On the other hand, this result supports the existing hypothesis about the beneficial effect of ACE inhibitors and ARBs (Ran et al., 2020a; Schiffrin et al., 2020) (standardly used in the treatment of hypertension). Similarly unintuitive correlation we report in the case of chronic diseases that are known to be relevant for the COVID-19 outcome. Namely, our data show, at very high statistical significance, a strong negative correlation of R_0 with the risk of death from a batch of chronic diseases (cardiovascular disease, cancer, diabetes, and chronic respiratory disease), agreeing in this regard with some recent research

(Chiang et al., 2020; Li et al., 2020). These diseases are identified as relevant comorbidities in the context of COVID-19, leading to a huge increase in the severity of the infection and poorer prognosis (An et al., 2020; Zheng et al., 2020) and, therefore, the discovered negative correlation comes as a surprise—particularly when contrasted to the positive correlation of obesity (where both are recognized risk factors in COVID-19 illness). One possible explanation is that the correlation may be due to potentially lower mobility of people with chronic diseases compared to the general mobility of the population. Additionally, it is possible that these people, being aware to belong to a high-risk group, behaved more cautiously even before the official introduction of social distancing measures.

According to our analysis, the prevalence of certain health-hazard habits is also significantly correlated to COVID-19 transmissibility. Chronic excessive alcohol consumption has, in general, a detrimental effect on immunity to viral and bacterial infections, which, judging by the strong positive correlation we obtained, most likely applies also to SARS-CoV-2 virus infection. This correlation contradicts the belief that alcohol can be used as a protective nostrum against COVID-19, which has spread in some countries and even led to cases of alcohol poisoning (Chick, 2020).

Regarding the impact of smoking on SARS-CoV-2 virus infection—the results are controversial (Chatkin and Godoy, 2020). The positive correlation of smoking with COVID-19 transmissibility that we obtained seems to support the reasoning that, since the SARS-CoV-2 virus enters cells by binding to angiotensin-converting enzyme 2 (ACE2) receptors and that the number of these receptors is significantly higher in the lungs of smokers, the smokers will be more affected and easily infected (Brake et al., 2020; Hoffmann et al., 2020). Accordingly, our result contradicts the hypothesis that a weakened immune response of smokers to virus infection may prove beneficial in the context of inflammation caused by intense cytokine release (Garufi et al., 2020).

Another result that addresses the association of unhealthy lifestyle with greater susceptibility to SARS-CoV-2 infection is the slight positive correlation we obtained for the prevalence of insufficient physical activity (IN) in adults, which is however not statistically significant. In this sense, in the case of COVID-19, we could not fully confirm the findings from (Jurak et al., 2020), who found that physical activity significantly reduces the risk of viral infections.

Despite the recent media interest (Gallagher, 2020), our findings neither could confirm that BCG immunization has any beneficial effect in the case of COVID-19, at least as far as reducing the risk of contracting and transmitting the disease is concerned. While it is known that the BCG vaccine provides some protection against various infectious agents, unfortunately, there is no clear evidence for such an effect against SARS-CoV-2 (O'Neill and Netea, 2020). Our analysis suggests that BCG immunization simply does not correlate with SARS-CoV-2 virus transmission.

SARS-CoV-2 target cells are typically capable of synthesizing ABH antigens and certain arguments exist, both theoretical and experimental, for a potential relation of blood groups

with COVID-19 progression and transmission (Guillon et al., 2008; Dai, 2020; Gérard et al., 2020). While the results of epidemiological studies on COVID-19 patients mostly support the proposed effect of blood groups on the development of COVID-19 disease, the relationship between virus transmission and blood group prevalence and Rh phenotype has been significantly less studied. Our analysis showed strong positive correlations of virus transmission with the presence of A blood group and Rh$^-$ phenotype, as well as strong negative correlations for B blood group and Rh$^+$ phenotype, while for AB and O blood group no significant correlations were obtained (**Figure 5C**). This result coincides significantly with the correlations obtained in a study conducted for 86 countries (Ansari-Lari and Saadat, 2020). However, another study focused on hospitalized patients in Turkey reported that the Rh$^+$ phenotype represents a predisposition to infection (Arac et al., 2020), contradicting our findings. Similar results regarding the Rh factor were obtained in a study (Latz et al., 2020) on hospitalized patients in the US (this study further reported no correlation of blood types with the severity of the disease). One way to reconcile these results with ours would be to speculate that the virus is more efficiently transmitted in a population with a higher proportion of Rh$^-$ phenotype because these people show a milder clinical picture compared to Rh$^+$, so their movement is not equally limited, which is why they have more ability to pass on the infection.

Our data (**Figure 6A**) shows a strong negative correlation with the date of the epidemic onset. Curiously, it seems that the later the epidemic started in a given country, it is more likely that the disease expansion will be slower. Instead of interpreting this result as an indication that the virus has mutated and changed its properties over such a short period, we offer the following simpler explanation: pandemic reached first those countries that are most interconnected with the rest of the world (at the same time, those are the countries characterized by great mobility of people overall), so it is expected that also the progression of the local epidemics in these countries is more rapid. Another contributing factor could be the effect of media, which had more time to raise awareness about the risks of COVID-19 in the countries that were hit later (Khajanchi et al., 2020b).

Another segment of our interest were air pollutants, shown in **Figure 6B**. Air pollution can have a detrimental effect on the human immune system and lead to the development (or to worsening) of respiratory diseases, including those caused by respiratory viral infections (Becker and Soukup, 1999; Copat et al., 2020). Several papers have already investigated air pollution in the context of COVID-19 and reported a positive correlation between the death rate due to COVID-19 and the concentration of PM2.5 in the environment (Wu et al., 2020; Yao et al., 2020c). Positive correlations were also found for the spread of the SARS-CoV-2 virus, but mainly by considering daily numbers of newly discovered cases—a method that, as we have already argued, may strongly depend on testing policies, as well as on state measures to combat the epidemic (Copat et al., 2020). It has been suggested that virus RNA can be adsorbed to airborne particles facilitating thus its spread over greater distances (Coccia, 2020; Setti et al., 2020), but these arguments were contested by examination of air samples in Wuhan (Contini and Costabile, 2020; Liu et al., 2020).

The latter conclusions concur with the results of a study in which no correlation was obtained between the basic reproductive number of SARS-CoV-2 infection for 154 Chinese cities and the concentration of PM2.5 and PM10 particles, while the correlation of these factors with the death rate (CFR) was shown (Ran et al., 2020b). The statistically insignificant and relatively weak correlations we obtained for PM2.5 and PM10 pollutants also do not support the hypothesis of a potentially significant role of these particles in the transmission of this virus. In contrast, significant positive correlations were shown by our analysis for concentrations of NO_2, SO_2, and CO in the air (although the correlation for CO is not statistically significant), which is generally supported by the results of other studies. For example, a positive correlation of NO_2 levels with the basic reproductive number of infection was obtained from data for 63 Chinese cities (Yao et al., 2020a). Also, it has been shown that the number of detected cases of COVID-19 in China is strongly positively correlated with the level of CO, while in Italy and the USA such correlation exists with NO_2 (Pansini and Fornacca, 2020). The mentioned study failed to establish a clear correlation with the level of SO_2. Possible mechanisms of interaction were also proposed (Daraei et al., 2020). Also, it is important to emphasize that the atmospheric concentration of NO_2 strongly depends on the levels of local exhaust emissions, so its correlation with virus transmission can be interpreted by the connection with the urban environment, characterized by more intensive traffic (Goldberg et al., 2020).

Finally, we have also obtained some interesting correlations of the meteorological parameters with R_0, shown in **Figures 6C,D**. The statistically very highly significant negative correlation of the basic reproductive number of SARS-CoV-2 virus infection with both the mean temperature and humidity obtained in our research (**Figure 6D**) is consistent with the results of other relevant papers, e.g., (Mecenas et al., 2020). For example, a similar correlation was obtained in a study that analyzed COVID-19 outbreak in the cities of Chile—a country that covers several climate zones, but where it is still safe to assume that social patterns of behavior and introduced epidemic control measures do not drastically differ throughout the country (Correa-Araneda et al., 2020). Effectively the same conclusion—that fewer COVID-19 cases were reported in countries with higher temperatures and humidities—was reached in a study covering over 200 countries in the world (Iqbal et al., 2020). While an established correlation between virus transmission and a certain factor is not, in general, a telltale sign of a direct causal relationship between them, in the case of temperature and humidity such connection is firmly indicated also by results of experimental research (Lowen et al., 2007; Casanova et al., 2010; Chan et al., 2011; van Doremalen et al., 2020). Nevertheless, some studies yielded different conclusions, most likely due to the method of calculating R_0 or due to choosing a small/uninformative sample of populations in which the number of infected cases was monitored (Guo et al., 2020; Lin et al., 2020; Yao et al., 2020b). For example, a study focused on the suburbs of New York, Queens, obtained a positive correlation between virus transmission and temperature, which seems unexpected given the prevailing observations of other studies (Adhikari and Yin,

2020). This result is most likely a consequence of analyzing data for a small area (Queens only) where the temperature varies in a relatively narrow range of values, as well as correlating the number of detected cases, which may be sensitive to variations in the testing procedure.

Another environmental agent that can destroy or inactivate viruses is UV radiation from sunlight, and the properties of a particular virus determine how long it can remain infectious when exposed to radiation. For example, epidemics of influenza have a seasonal character precisely due to the susceptibility of influenza viruses to UV radiation (Sagripanti and Lytle, 2007). Our analysis found, at very high statistical significance, a strong negative correlation between the transmission of the SARS-CoV-2 virus and the intensity of UV radiation, which is consistent with the results of other studies obtained for the cities of Brazil and the provinces of Iran (Ahmadi et al., 2020; Mendonça et al., 2020). It is worth mentioning that lower temperatures, humidity, and sunlight levels usually occur in combination and directly affect not only the virus but also the human behavior, so the observed higher transmission of the virus in such conditions can alternatively be interpreted by indirect effects of other factors that act together in cold weather, such as more time spent indoors where the virus spreads more easily, or weakening of the immune system that increases susceptibility to infections (Abdullahi et al., 2020).

While the results related to COVID-19 correlations with temperature, humidity, and UV radiation are fairly frequent in the literature, this is less so for the results on the precipitation levels. Very few other studies have examined the association of precipitation with SARS-CoV-2 transmission, with either no correlation found (Pourghasemi et al., 2020), or looking at precipitation as a surrogate for humidity and generally receiving a negative correlation with infection rate (Araujo and Naimi, 2020; Coro, 2020). Our results, however, shown in **Figure 6D**, confirm natural expectations: just like humidity, the precipitation exhibits a strong negative correlation with R_0, only slightly lower than in the case of T, H, and UV, at a very high level of statistical significance. Such results also concur with some general conclusions about the behavior of similar viruses (Agrawal et al., 2009; Pica and Bouvier, 2012).

Our analysis did not reveal any statistically significant correlation either between the wind speed or between air pressure and SARS-CoV-2 transmissibility. In the case of wind speed, this result agrees with the findings in some other papers (Gupta et al., 2020; Oliveiros et al., 2020). A positive correlation of wind speed with COVID-19 transmissibility was obtained in a study in Chilean cities, but, as the authors themselves note, the interpretation of the effect of this factor is complicated by its observed significant interaction with temperature (Correa-Araneda et al., 2020). The role of wind in transmitting the virus to neighboring buildings is predicted by the SARS virus spread model within the Amoy Gardens residential complex in Hong Kong, but such an effect may relate to local air currents and virus transmission over relatively short distances and does not imply a correlation of mean wind speeds in the area with virus transmission (McKinney et al., 2006; Pica and Bouvier, 2012). As for the air pressure, the potential connection is hardly at all investigated in the literature. An exception is a study (Cambaza et al., 2020) reporting a positive correlation of air pressure with the number of COVID-19 cases in parts of Mozambique, but our results do not confirm such a conclusion.

CONCLUSION

While there is by now a significant amount of research on a crucial problem of how environmental factors affect COVID19 spread, several features set this analysis apart from the existing research. First is the applied methodology: instead of basing analysis directly on the number of detected COVID-19 cases (or some of its simple derivatives), we employ an adapted SEIR model to extract the basic reproduction number R_0 from the initial stage of the epidemic. By taking into account only data in the exponential growth regime, i.e., before the social measures took effect (as explained in the "Methods" section), we ensured that the correlations we have later identified were not confounded with the effects of local COVID-19 policies. Even more importantly, our method is also invariant to variations in COVID-19 testing practices, which, as is well known, used to vary in quite an unpredictable manner between different countries. Another important factor is the large geographical scope of our research: we collected data from 118 countries worldwide, more precisely, from all the countries that were above a certain threshold for the number of confirmed COVID-19 cases (except for several countries with clearly irregular early growth data). The third factor was the number of analyzed parameters: we calculated correlations for the selected 42 different variables (of more than a hundred that we initially considered overall) and looked for viable interpretations of the obtained results.

These results should also help in resolving some of the existing disputes in the literature. For example, our findings indicate that correlation of HDI with R_0 is not a consequence of the COVID-19 testing bias, as was occasionally argued. Of the opposing opinions, our data seem to support assertions that blood types are indeed related to COVID-19 transmissibility, as well as arguments that the higher prevalence of smoking does increase the virus transmissibility (though weakly). On the other hand, in the dispute about the effects of the pollution, our correlations give an edge to claims that there is no correlation between PM2.5 and PM10 particles and transmissibility (whereas we agree with the prevailing conclusions about the positive correlation of other considered pollutants). In the case of the effects of the wind, based on the obtained results we tend to side with those denying any connection. In certain cases our findings contradict popular narratives: there are no clear indications that either number of refugees or physical inactivity intensifies the spread of COVID-19. Unfortunately, our data also suggest that BCG immunization may not help in subduing the epidemic. Additionally, the obtained correlations hint to possible new alleys of research, e.g., those that would help us understand the connection between cholesterol levels and SARS-CoV-2 transmissibility.

Overall, we believe that the presented results can be a useful contribution to the ongoing attempts to better understand the first pandemic of the twenty-first century—and the better we understand it, the sooner we may hope to overcome it.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MarD and MagD conceived the research. The work was supervised by IS, MagD, and MarD. Data acquisition by OM, IS, DZ, and AR. Data analysis by DZ, OM, and IS, with the help of MagD. Figures and tables made by DZ and AR with the help of MagD. A literature search by AR. Manuscript written by IS, AR, and MarD. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2020.617841/full#supplementary-material

REFERENCES

Abdullahi, I. N., Emeribe, A. U., Mustapha, J. O., Fasogbon, S. A., Ofor, I. B., Opeyemi, I. S., et al. (2020). Exploring the genetics, ecology of SARS-COV-2 and climatic factors as possible control strategies against COVID-19. *Infez. Med.* 28, 166–173.

Adhikari, A., and Yin, J. (2020). Short-term effects of ambient ozone, PM2.5, and meteorological factors on COVID-19 confirmed cases and deaths in queens, New York. *Int. J. Environ. Res. Public Health* 17:4047. doi: 10.3390/ijerph17114047

Agrawal, A. S., Sarkar, M., Chakrabarti, S., Rajendran, K., Kaur, H., Mishra, A. C., et al. (2009). Comparative evaluation of real-time PCR and conventional RT-PCR during a 2 year surveillance for influenza and respiratory syncytial virus among children with acute respiratory infections in Kolkata, India, reveals a distinct seasonality of infection. *J. Med. Microbiol.* 58, 1616–1622. doi: 10.1099/jmm.0.011304-0

Ahmadi, M., Sharifi, A., Dorosti, S., Ghoushchi, S. J., and Ghanbari, N. (2020). Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Sci. Total Environ.* 729:138705. doi: 10.1016/j.scitotenv.2020. 138705

An, C., Lim, H., Kim, D. W., Chang, J. H., Choi, Y. J., and Kim, S. W. (2020). Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci. Rep.* 10:18716. doi: 10.1038/s41598-020-75767-2

Anderson, R. M., and May, R. M. (1992). *Infectious Diseases of Humans: Dynamics and Control.* New York, NY:: Oxford University Press.

Ansari-Lari, M., and Saadat, M. (2020). The morbidity and mortality of COVID-19 are associated with ABO and Rh blood groups. *Eur. J. Prev. Cardiol.* doi: 10.1177/2047487320939216. [Epub ahead of print].

Arac, E., Solmaz, I., Akkoc, H., Donmezdil, S., Karahan, Z., Kaya, S., et al. (2020). Association between the Rh blood group and the Covid-19 susceptibility. *UHOD-Uluslar. Hematol.* 30, 81–86. doi: 10.4999/uhod.204247

Araujo, M. B., and Naimi, B. (2020). Spread of SARS-CoV-2 coronavirus likely constrained by climate. medRxiv [Preprint]. Available online at https://europepmc.org/article/ppr/ppr117563 (accessed December 16, 2020).

Bar-On, Y. M., Flamholz, A., Phillips, R., and Milo, R. (2020). SARS-CoV-2 (COVID-19) by the numbers. *Elife* 9:e57309. doi: 10.7554/eLife.57309.sa2

Becker, S., and Soukup, J. M. (1999). Exposure to urban air particulates alters the macrophage-mediated inflammatory response to respiratory viral infection. *J. Toxicol. Environ. Health Part A* 57, 445–457. doi: 10.1080/009841099157539

Berg, M. K., Yu, Q., Salvador, C. E., Melani, I., and Kitayama, S. (2020). Mandated Bacillus Calmette-Guérin (BCG) vaccination predicts flattened curves for the spread of COVID-19. *Sci. Adv.* 6:eabc1463. doi: 10.1126/sciadv.abc1463

Brake, S. J., Barnsley, K., Lu, W., McAlinden, K. D., Eapen, M. S., and Sohal, S. S. (2020). Smoking upregulates angiotensin-converting enzyme-2 receptor: a potential adhesion site for novel coronavirus SARS-CoV-2 (Covid-19). *J. Clin. Med.* 9:841. doi: 10.3390/jcm9030841

Cambaza, E. M., Viegas, G. C., and Cambaza, C. M. (2020). Potential impact of temperature and atmospheric pressure on the number of cases of COVID-19 in Mozambique, Southern Africa. *J. Public Health Epidemiol.* 12, 246–260. doi: 10.5897/JPHE2020.1258

Carozzi, F. (2020). Urban density and COVID-19. *Institute for the Study of Labor (IZA)* [Online], 13440. Available online at https://ssrn.com/abstract=3643204 (accessed December 16, 2020).

Casanova, L. M., Jeon, S., Rutala, W. A., Weber, D. J., and Sobsey, M. D. (2010). Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl. Environ. Microbiol.* 76, 2712–2717. doi: 10.1128/AEM.02291-09

Chan, K.-H., Peiris, J. M., Lam, S., Poon, L., Yuen, K., and Seto, W. H. (2011). The effects of temperature and relative humidity on the viability of the SARS coronavirus. *Adv. Virol.* 2011:734690. doi: 10.1155/2011/734690

Chatkin, J. M., and Godoy, I. (2020). Are smoking, environmental pollution, and weather conditions risk factors for COVID-19? *J. Bras. Pneumol.* 46:e20200183. doi: 10.36416/1806-3756/e20200183

Chiang, W.-H., Liu, X., and Mohler, G. (2020). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *medRxiv [Preprint].* Available online at: https://europepmc.org/article/ppr/ppr172763 (accessed December 16, 2020).

Chick, J. (2020). Alcohol and COVID-19. *Alcohol Alcohol.* 55, 341–342. doi: 10.1093/alcalc/agaa039

Coccia, M. (2020). The effects of atmospheric stability with low wind speed and of air pollution on the accelerated transmission dynamics of COVID-19. *Int. J. Environ. Stud.* doi: 10.1080/00207233.2020.1802937. [Epub ahead of print].

Cohen, J., and Kupferschmidt, K. (2020). Countries test tactics in 'war' against COVID-19. *Science* 367, 1287–1288. doi: 10.1126/science.367.6484.1287

Contini, D., and Costabile, F. (2020). Does air pollution influence COVID-19 outbreaks? *Atmosphere* 11:377. doi: 10.3390/atmos11040377

Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., et al. (2020). The role of air pollution (PM and NO$_2$) in COVID-19 spread and lethality: a systematic review. *Environ. Res.* 191:110129. doi: 10.1016/j.envres.2020.110129

Copiello, S., and Grillenzoni, C. (2020). The spread of 2019-nCoV in China was primarily driven by population density. Comment on "Association between short-term exposure to air pollution and COVID-19 infection: evidence from China" by Zhu et al. *Sci. Total Environ.* 744:141028. doi: 10.1016/j.scitotenv.2020.141028

Coro, G. (2020). A global-scale ecological niche model to predict SARS-CoV-2 coronavirus infection rate. *Ecol. Modell.* 431:109187. doi: 10.1016/j.ecolmodel.2020.109187

Correa-Araneda, F., Ulloa-Yañez, A., Núñez, D., Boyero, L., Tonin, A., Cornejo, A., et al. (2020). Environmental determinants of COVID-19 transmission across a

wide climatic gradient in Chile. Research Square [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr166276 (accessed December 16, 2020).

Dai, X. (2020). ABO blood group predisposes to COVID-19 severity and cardiovascular diseases. *Eur. J. Prev. Cardiol.* 27, 1436–1437. doi: 10.1177/2047487320922370

Daraei, H., Toolabian, K., Kazempour, M., and Javanbakht, M. (2020). The role of the environment and its pollution in the prevalence of COVID-19. *J. Infect.* 81, e168–e169. doi: 10.1016/j.jinf.2020.06.019

Day, M. (2020). Covid-19: four fifths of cases are asymptomatic, China figures indicate. *BMJ* 369:m1375. doi: 10.1136/bmj.m1375

Djordjevic, M., Djordjevic, M., Salom, I., Rodic, A., Zigic, D., Milicevic, O., et al. (2020). COVID-19 puzzle in China: a serendipitous interplay between transmissibility and social distancing measures. arXiv [Preprint]. Available online at: https://arxiv.org/abs/2005.09630 (accessed December 16, 2020).

Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *Lancet. Infect. Dis.* 20, 533–534. doi: 10.1016/S1473-3099(20)30120-1

Eubank, S., Eckstrand, I., Lewis, B., Venkatramanan, S., Marathe, M., and Barrett, C. (2020). Commentary on Ferguson, et al., "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand". *Bull. Math. Biol.* 82:52. doi: 10.1007/s11538-020-00726-x

European Commission Global Human Settlement (2020). *Urban Centre Database UCDB R2019A* [Online]. Available online at: https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php (accessed September 26, 2020).

Fan, C., Cai, T., Gai, Z., and Wu, Y. (2020). The relationship between the migrant population's migration network and the risk of COVID-19 transmission in China—empirical analysis and prediction in prefecture-level cities. *Int. J. Environ. Res. Public Health* 17:2630. doi: 10.3390/ijerph17082630

Fareed, Z., Iqbal, N., Shahzad, F., Shah, S. G. M., Zulfiqar, B., Shahzad, K., et al. (2020). Co-variance nexus between COVID-19 mortality, humidity, and air quality index in Wuhan, China: new insights from partial and multiple wavelet coherence. *Air Qual. Atmos. Health* 8, 1–10. doi: 10.1007/s11869-020-00847-1

Gallagher, J. (2020, October 11). BCG: can a vaccine from 1921 save lives from Covid-19? *BBC News*.

Gangemi, S., Billeci, L., and Tonacci, A. (2020). Rich at risk: socio-economic drivers of COVID-19 pandemic spread. *Clin. Mol. Allergy* 18:12. doi: 10.1186/s12948-020-00127-4

Garufi, G., Carbognin, L., Orlandi, A., Tortora, G., and Bria, E. (2020). Smoking habit and hospitalization for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-related pneumonia: the unsolved paradox behind the evidence. *Eur. J. Intern. Med.* 77, 121–122. doi: 10.1016/j.ejim.2020.04.042

Gérard, C., Maggipinto, G., and Minon, J.-M. (2020). COVID-19 and ABO blood group: another viewpoint. *Brit. J. Haematol.* 190, e57–e94. doi: 10.1111/bjh.16884

Global Data Lab (2020). *Subnational Human Development Index (4.0)* [Online]. Available online at://globaldatalab.org/shdi/(accessed September 27, 2020).

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., and Streets, D. G. (2020). Disentangling the impact of the COVID-19 lockdowns on urban NO2 from natural variability. *Geophys. Res. Lett.* 47:e2020GL089269. doi: 10.1002/essoar.10503396.1

Guillon, P., Clément, M., Sébille, V., Rivain, J.-G., Chou, C.-F., Ruvoën-Clouet, N., et al. (2008). Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies. *Glycobiology* 18, 1085–1093. doi: 10.1093/glycob/cwn093

Guo, X.-J., Zhang, H., and Zeng, Y.-P. (2020). Transmissibility of COVID-19 in 11 major cities in China and its association with temperature and humidity in Beijing, Shanghai, Guangzhou, and Chengdu. *Infect. Dis. Poverty* 9:87. doi: 10.1186/s40249-020-00708-0

Gupta, A., Banerjee, S., and Das, S. (2020). Significance of geographical factors to the COVID-19 outbreak in India. *Model. Earth Syst. Environ.* 6, 2645–2653. doi: 10.1007/s40808-020-00838-2

Haidar, Y. M., and Cosman, B. C. (2011). Obesity epidemiology. *Clin. Colon Rectal Surg.* 24, 205–210. doi: 10.1055/s-0031-1295684

Hargreaves, S., Kumar, B. N., McKee, M., Jones, L., and Veizis, A. (2020). Europe's migrant containment policies threaten the response to covid-19. *BMJ* 368:m1213. doi: 10.1136/bmj.m1213

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280. doi: 10.1016/j.cell.2020.02.052

Iqbal, M. M., Abid, I., Hussain, S., Shahzad, N., Waqas, M. S., and Iqbal, M. J. (2020). The effects of regional climatic condition on the spread of COVID-19 at global scale. *Sci. Total Environ.* 739:140101. doi: 10.1016/j.scitotenv.2020.140101

Jordan, R. E., Adab, P., and Cheng, K. K. (2020). Covid-19: risk factors for severe disease and death. *BMJ* 368:m1198. doi: 10.1136/bmj.m1198

Jurak, G., Morrison, S. A., Leskošek, B., Kovač, M., HadŽić, V., Vodičar, J., et al. (2020). Physical activity recommendations during the coronavirus disease-2019 virus outbreak. *J. Sport Health Sci.* 9, 325–327. doi: 10.1016/j.jshs.2020.05.003

Keeling, M. J., and Rohani, P. (2011). *Modeling Infectious Diseases in Humans and Animals.* Princeton, NJ: Princeton University Press. doi: 10.2307/j.ctvcm4gk0

Khajanchi, S., Bera, S., and Roy, T. K. (2020a). Mathematical analysis of the global dynamics of a HTLV-I infection model, considering the role of cytotoxic T-lymphocytes. *Math. Comput. Simulat.* 180, 354–378. doi: 10.1016/j.matcom.2020.09.009

Khajanchi, S., and Sarkar, K. (2020). Forecasting the daily and cumulative number of cases for the COVID-19 pandemic in India. *Chaos* 30:071101. doi: 10.1063/5.0016240

Khajanchi, S., Sarkar, K., Mondal, J., and Perc, M. (2020b). Dynamics of the COVID-19 pandemic in India. arXiv [Preprint]. Available at: https://arxiv.org/abs/2005.06286 (accessed December 16, 2020). doi: 10.21203/rs.3.rs-27112/v1

Kimball, A., Hatfield, K. M., Arons, M., James, A., Taylor, J., Spicer, K., et al. (2020). Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020. *MMWR Morb. Mortal. Wkly Rep.* 69, 377–381. doi: 10.15585/mmwr.mm6913e1

Latz, C. A., DeCarlo, C., Boitano, L., Png, C. M., Patell, R., Conrad, M. F., et al. (2020). Blood type and outcomes in patients with COVID-19. *Ann. Hematol.* 99, 2113–2118. doi: 10.1007/s00277-020-04169-1

Li, M., Zhang, Z., Cao, W., Liu, Y., Du, B., Chen, C., et al. (2020). Identifying novel factors associated with COVID-19 transmission and fatality using the machine learning approach. *Sci. Total Environ.* 142810. doi: 10.1016/j.scitotenv.2020.142810. [Epub ahead of print].

Lin, S., Wei, D., Sun, Y., Chen, K., Yang, L., Liu, B., et al. (2020). Region-specific air pollutants and meteorological parameters influence COVID-19: a study from mainland China. *Ecotoxicol. Environ. Saf.* 204:111035. doi: 10.1016/j.ecoenv.2020.111035

Liu, Y., Ning, Z., Chen, Y., Guo, M., Gali, N. K., Sun, L., et al. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature* 582, 557–560. doi: 10.1038/s41586-020-2271-3

Lowen, A. C., Mubareka, S., Steel, J., and Palese, P. (2007). Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog.* 3, 1470–1476. doi: 10.1371/journal.ppat.0030151

Maier, B. F., and Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science* 368, 742–746. doi: 10.1126/science.abb4557

Martcheva, M. (2015). *An Introduction to Mathematical Epidemiology.* Boston, MA: Springer. doi: 10.1007/978-1-4899-7612-3

Maslov, S., and Goldenfeld, N. (2020). Window of opportunity for mitigation to prevent overflow of ICU capacity in Chicago by COVID-19. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr118721 (accessed December 16, 2020).

McKinney, K. R., Yu, Y. G., and Lewis, T. G. (2006). Environmental transmission of SARS at amoy gardens. *J. Environ. Health* 68, 26–30.

Mecenas, P., Bastos, R., Vallinoto, A. C. R., and Normando, D. (2020). Effects of temperature and humidity on the spread of COVID-19: a systematic review. *PLoS ONE* 15:e0238339. doi: 10.1371/journal.pone.0238339

Mendonça, F., Anjos, M., Collischonn, E., Murara, P., DE, F., Limberger, L., et al. (2020). Climate and Covid-19–upgrade and solar radiation influences based on Brazil cases. Research Square [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr174953 (accessed December 16, 2020).

Najafimehr, H., Mohamed Ali, K., Safari, S., Yousefifard, M., and Hosseini, M. (2020). Estimation of basic reproduction number for COVID-19 and the reasons for its differences. *Int. J. Clin. Pract.* 74:e13518. doi: 10.1111/ijcp.13518

NASA Langley Research Center (2020). *The Prediction of Worldwide Energy Resources (POWER) Project [Online].* Available online at: https://power.larc.nasa.gov/ (accessed October 6, 2020).

Oliveiros, B., Caramelo, L., Ferreira, N. C., and Caramelo, F. (2020). Role of temperature and humidity in the modulation of the doubling time of COVID-19 cases. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr116498 (accessed December 16, 2020).

O'Neill, L. A., and Netea, M. G. (2020). BCG-induced trained immunity: can it offer protection against COVID-19? *Nat. Rev. Immunol.* 20, 335–337. doi: 10.1038/s41577-020-0337-y

Our World in Data (2020). *Research and Data to Make Progress Against the World's Largest Problems* [Online]. Available online at: https://ourworldindata.org/ (accessed May, 2020).

Pansini, R., and Fornacca, D. (2020). COVID-19 higher induced mortality in Chinese regions with lower air quality. medRxiv [Preprint]. Available online at //www.medrxiv.org/content/10.1101/2020.05.28.20115832v2 (accessed December 16, 2020).

Pawelec, G., and Larbi, A. (2008). Immunity and ageing in man: annual review 2006/2007. *Exp. Gerontol.* 43, 34–38. doi: 10.1016/j.exger.2007.09.009

Perkins, T. A., and España, G. (2020). Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions. *Bull. Math. Biol.* 82:118. doi: 10.1007/s11538-020-00795-y

Pica, N., and Bouvier, N. M. (2012). Environmental factors affecting the transmission of respiratory viruses. *Curr. Opin. Virol.* 2, 90–95. doi: 10.1016/j.coviro.2011.12.003

Pourghasemi, H. R., Pouyan, S., Heidari, B., Farajzadeh, Z., Fallah Shamsi, S. R., Babaei, S., et al. (2020). Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *Int. J. Infect. Dis.* 98, 90–108. doi: 10.1016/j.ijid.2020.06.058

Pranata, R., Lim, M. A., Huang, I., Raharjo, S. B., and Lukito, A. A. (2020). Hypertension is associated with increased mortality and severity of disease in COVID-19 pneumonia: a systematic review, meta-analysis and meta-regression. *J. Renin Angiotensin Aldosterone Syst.* 21:1470320320926899. doi: 10.1177/1470320320926899

Qu, G., Li, X., Hu, L., and Jiang, G. (2020). An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19). *Environ. Sci. Technol.* 54, 3730–3732. doi: 10.1021/acs.est.0c01102

Radenkovic, D., Chawla, S., Pirro, M., Sahebkar, A., and Banach, M. (2020). Cholesterol in relation to COVID-19: should we care about it? *J. Clin. Med.* 9:1909. doi: 10.3390/jcm9061909

Ran, J., Song, Y., Zhuang, Z., Han, L., Zhao, S., Cao, P., et al. (2020a). Blood pressure control and adverse outcomes of COVID-19 infection in patients with concomitant hypertension in Wuhan, China. *Hypertens. Res.* 43, 1267–1276. doi: 10.1038/s41440-020-00541-w

Ran, J., Zhao, S., Han, L., Qiu, Y., Cao, P., Yang, Z., et al. (2020b). Effects of particulate matter exposure on the transmissibility and case fatality rate of COVID-19: a Nationwide Ecological Study in China. *J. Travel Med.* 27:taaa133. doi: 10.1093/jtm/taaa133

Rashed, E. A., Kodera, S., Gomez-Tames, J., and Hirata, A. (2020). Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: multi-prefecture study in Japan. *Int. J. Environ. Res. Public Health* 17:5354. doi: 10.3390/ijerph17155354

Ruiz, J. I., Nuhu, K., McDaniel, J. T., Popoff, F., Izcovich, A., and Criniti, J. M. (2015). Inequality as a powerful predictor of infant and maternal mortality around the world. *PLoS ONE* 10:e0140796. doi: 10.1371/journal.pone.0140796

Rychter, A. M., Zawada, A., Ratajczak, A. E., Dobrowolska, A., and Krela-Kazmierczak, I. (2020). Should patients with obesity be more afraid of COVID-19? *Obes. Rev.* 21:e13083. doi: 10.1111/obr.13083

Sagar, A. D., and Najam, A. (1998). The human development index: a critical review. *Ecol. Econ.* 25, 249–264. doi: 10.1016/S0921-8009(97)00168-7

Sagripanti, J.-L., and Lytle, C. D. (2007). Inactivation of influenza virus by solar radiation. *Photochem. Photobiol.* 83, 1278–1282. doi: 10.1111/j.1751-1097.2007.00177.x

Samui, P., Mondal, J., and Khajanchi, S. (2020). A mathematical model for COVID-19 transmission dynamics with a case study of India. *Chaos Solitons Fractals* 140:110173. doi: 10.1016/j.chaos.2020.110173

Sarkar, K., Khajanchi, S., and Nieto, J. J. (2020). Modeling and forecasting the COVID-19 pandemic in India. *Chaos Solitons Fractals* 139:110049. doi: 10.1016/j.chaos.2020.110049

Sattar, N., McInnes, I. B., and McMurray, J. J. (2020). Obesity a risk factor for severe COVID-19 infection: multiple potential mechanisms. *Circulation* 142, 4–6. doi: 10.1161/CIRCULATIONAHA.120.047659

Scheffer, M. (2009). *Critical Transitions in Nature and Society.* Princeton, NJ: Princeton University Press. doi: 10.1515/9781400833276

Schiffrin, E. L., Flack, J. M., Ito, S., Muntner, P., and Webb, R. C. (2020). Hypertension and COVID-19. *Am. J. Hypertens.* 33, 373–374. doi: 10.1093/ajh/hpaa057

Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., et al. (2020). SARS-Cov-2RNA found on particulate matter of Bergamo in Northern Italy: first evidence. *Environ. Res.* 188:109754. doi: 10.1016/j.envres.2020.109754

Shahid, Z., Kalayanamitra, R., McClafferty, B., Kepko, D., Ramgobin, D., Patel, R., et al. (2020). COVID-19 and older adults: what we know. *J. Am. Geriatr. Soc.* 68, 926–929. doi: 10.1111/jgs.16472

Singh, K., and Agarwal, A. (2020). Impact of weather indicators on the COVID-19 outbreak: a multi-state study in India. medRxiv [Preprint]. Available online at https://europepmc.org/article/ppr/ppr176027 (accessed December 16, 2020).

Thangriyal, S., Rastogi, A., Tomar, A., and Baweja, S. (2020). Impact of temperature and sunshine duration on daily new cases and death due to COVID-19. medRxiv [Preprint]. Available online at: https://europepmc.org/article/ppr/ppr176097 (accessed December 16, 2020).

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China, 2020. *China CDC Weekly* [Online], 2. Available online at: https://slma.lk/wp-content/uploads/2020/02/TheEpidemiologicalCharacteristicsofanOutbreakof2019NovelCoronavirusDiseases28COVID-1929E28094China2C20201.pdf. doi: 10.46234/ccdcw2020.032

Tian, H., Liu, Y., Li, Y., Wu, C. H., Chen, B., Kraemer, M. U. G., et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368, 638–642. doi: 10.1126/science.abb6105

Tosepu, R., Gunawan, J., Effendy, D. S., Ahmad, O. A. I., Lestari, H., Bahar, H., et al. (2020). Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Sci. Total Environ.* 725:138436. doi: 10.1016/j.scitotenv.2020.138436

van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* 382, 1564–1567. doi: 10.1056/NEJMc2004973

Ward, D. (2020). *Actions Speak Louder Than Age: Explaining Wide Variations in COVID-19 Deaths* [Online]. Available online at: https://www.researchgate.net/publication/341599695_Actions_Speak_Louder_than_Age_Explaining_Wide_Variations_in_COVID-19_Deaths (accessed December 16, 2020).

Weitz, J. S., Beckett, S. J., Coenen, A. R., Demory, D., Dominguez-Mirazo, M., Dushoff, J., et al. (2020). Modeling shield immunity to reduce COVID-19 epidemic spread. *Nat. Med.* 26, 849–854. doi: 10.1038/s41591-020-0895-3

World Bank (2020a). *Global Economic Prospects, June 2020.* Washington, DC: World Bank.

World Bank (2020b). *World Bank Open Data* [Online]. Available online at: https://www.worldbank.org/ (accessed May, 2020).

World Health Organization (2020). *The Global Health Observatory: Explore a World of Health Data* [Online]. Available online at: https://www.who.int/data/gho (accessed October 6, 2020).

World Weather Online (2020). *Local Weather History API* [Online]. Available online at: https://www.worldweatheronline.com/developer/api/historical-weather-api.aspx (accessed December 9, 2020).

Worldometer (2020). *COVID-19 Coronavirus Pandemic* [Online]. Available online at: https://www.worldometers.info/coronavirus/ (accessed June 14, 2020).

Wu, X., Nethery, R. C., Sabath, M. B., Braun, D., and 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. doi: 10.1126/sciadv.abd4049

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., et al. (2020a). Ambient nitrogen dioxide pollution and spread ability of COVID-19 in Chinese cities. *Ecotoxicol. Environ. Saf.* 208:111421. doi: 10.1016/j.ecoenv.2020.111421

Yao, Y., Pan, J., Liu, Z., Meng, X., Wang, W., Kan, H., et al. (2020b). No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. *Eur. Respir. J.* 55:2000517. doi: 10.1183/13993003.00517-2020

Yao, Y., Pan, J., Wang, W., Liu, Z., Kan, H., Qiu, Y., et al. (2020c). Association of particulate matter pollution and case fatality rate of COVID-19 in 49 Chinese cities. *Sci. Total Environ.* 741:140396. doi: 10.1016/j.scitotenv.2020.140396

Yu, Q., Salvador, C., Melani, I., Berg, M., Neblett, E., and Kitayama, S. (2020). Racial residential segregation and economic disparity jointly exacerbate the COVID-19 fatality in large American cities. *PsyArXiv [Preprint]*. Available online at: https://psyarxiv.com/xgbpy/ (accessed December 16, 2020).

Zheng, Z., Peng, F., Xu, B., Zhao, J., Liu, H., Peng, J., et al. (2020). Risk factors of critical and mortal COVID-19 cases: a systematic literature review and meta-analysis. *J. Infect.* 81:e16–e25. doi: 10.1016/j.jinf.2020.04.021