

RESEARCH

Open Access

Predicting the number of COVID-19 infections and deaths in USA



Amarachukwu Felix Ebubeogu^{1*†} , Chamberline Ekene Ozigbu², Kholoud Maswadi³, Azizi Seixas⁴, Paulinus Ofem⁵ and Donaldson F. Conserve⁶

Abstract

Background: Uncertainties surrounding the 2019 novel coronavirus (COVID-19) remain a major global health challenge and requires attention. Researchers and medical experts have made remarkable efforts to reduce the number of cases and prevent future outbreaks through vaccines and other measures. However, there is little evidence on how severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection entropy can be applied in predicting the possible number of infections and deaths. In addition, more studies on how the COVID-19 infection density contributes to the rise in infections are needed. This study demonstrates how the SARS-CoV-2 daily infection entropy can be applied in predicting the number of infections within a given period. In addition, the infection density within a given population attributes to an increase in the number of COVID-19 cases and, consequently, the new variants.

Results: Using the COVID-19 initial data reported by Johns Hopkins University, World Health Organization (WHO) and Global Initiative on Sharing All Influenza Data (GISAID), the result shows that the original SAR-CoV-2 strain has $R_0 < 1$ with an initial infection growth rate entropy of 9.11 *bits* for the United States (U.S.). At close proximity, the average infection time for an infected individual to infect others within a susceptible population is approximately 7 minutes. Assuming no vaccines were available, in the U.S., the number of infections could range between 41,220,199 and 82,440,398 in late March 2022 with approximately, 1,211,036 deaths. However, with the available vaccines, nearly 48 Million COVID-19 cases and 706, 437 deaths have been prevented.

Conclusion: The proposed technique will contribute to the ongoing investigation of the COVID-19 pandemic and a blueprint to address the uncertainties surrounding the pandemic.

Keywords: 2019 novel coronavirus, COVID-19, SARS-CoV-2, Infection entropy, Infection density

Introduction

The COVID-19 outbreak has remained a universal health concern that requires urgent attention. Notably, Scientists have pointed out that COVID-19 is the newest coronavirus species with public health emergency [1–3]. Although several studies have been carried out aimed to provide useful information regarding the coronavirus pandemic [4–7], studies are still ongoing to uncover the

root cause of the pandemic as well as solutions to address the outbreak. In December 2019, there were clusters of pneumonia cases in China. Later, investigations discovered that an unknown virus caused such clusters of pneumonia [8, 9]. The unknown virus is currently called the 2019 novel coronavirus. Coronaviruses are a large group of viruses that consist of core genetic materials surrounded by specific protein spikes [10, 11]. Despite their unique nature, there are different types of coronaviruses that cause respiratory symptoms. These symptoms may range from the common cold to pneumonia, as in China's case, where it was first identified. These symptoms may be mild in most cases, whereas some cases are severe. For instance, fever, cough, and shortness of breath may

*Correspondence: felixbosken@hotmail.com

[†]Amarachukwu Felix Ebubeogu contributed equally to this work.

¹Department of Software Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

Full list of author information is available at the end of the article



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

be signs of mild symptoms. On the other hand, pneumonia, kidney failure, and death characterize severe cases. However, some kinds of coronaviruses are responsible for severe cases, such as severe acute respiratory syndrome coronavirus (SARS-CoV), first discovered in China in 2002–2003 [12–14]. Another type of coronavirus that can also cause severe health damage is the Middle East respiratory syndrome-related coronavirus (MERS-CoV), identified in the Kingdom of Saudi Arabia in 2012 [15, 16].

Despite the symptoms of MERS-CoV, which include sore throat, headache, fever, mild cough, tiredness, runny nose and diarrhea, its transmission among humans appears to be less harmful [17]. SARS-CoV-2 is characterized by more contagious variants [18–21]. In terms of structure, the S proteins of SARS-CoV and SARS-CoV-2 are similar [22, 23]. In terms of R_0 , a daily reproduction number of 2.68 was reported by Wu et al. for SARS-CoV-2 [22], equivalent to the reports by both the WHO and the Chinese Center for Disease Control [24, 25]. A previous study reported an actual R_0 value between 2.0 and 2.5 for SARS-CoV-2, which remains disputed [26]. However, the disputed R_0 values for SARS-CoV-2 are lower than the 1.7 and 1.9 R_0 for SARS and $R_0 < 1$ for MERS, respectively, [26]. In addition, the R_0 values for SARS-CoV-2 have been estimated to range between 2.24 and 3.58 [27], while another study reported a range between 2.0 and 5.0 for SARS [28]. Similarly, certain predictive models have suggested R_0 values of 3.8 (95% CI, 3.6–4.0) [29] and 3.11 (95% CI, 2.39–4.13) for SARS-CoV-2 [30]. For SARS, the R_0 was estimated to be approximately 3.0 if adequate control measures were not in place [31]. A high value of 5.8 (confidence interval: 4.7–7.3) for SARS-CoV-2 was reported for the U.S., and a range between 3.6 and 6.1 was reported for some countries in Europe [32]. Notably, a group of researchers reported a higher R_0 value of 6.47 for SARS-CoV-2 [33]. These high values of R_0 indicate that the SARS-CoV-2 virus has the ability to rapidly mutate and spread [2, 34].

Notably, COVID-19 poses a health threat and an economic threat across the globe [35]. Obviously, the gross domestic product (GDP) of almost all the affected countries has dropped tremendously, and as such, goods and services are affected along the supply chains. Besides, millions of schools from kindergarten to institutions of higher learning remain closed. Millions of both private and public companies, as well as their respective employers and employees, are under lockdown. Consequently, millions of workers have lost their jobs as a result of this pandemic.

Notwithstanding the remarkable achievements of the existing studies on providing useful information regarding the COVID-19 pandemic, the following gaps exist in the literature: (1) how the daily infection density attributes to increase in number of cases, (2) how daily infection entropy can be applied in predicting the number of cases

and deaths and (3) the average time it may take for an infected individual to infect a susceptible population in close proximity.

Therefore, more studies are needed to support medical expert investigations as well as in their decision-making processes to uncover novel preventive measures to complement the available vaccines for the 2019 novel coronavirus. Hence, to better understand and characterize the initial behaviour of the virus, the current study aggregates the number of the initial reported COVID-19 cases before the emergence of the new variants and the corresponding numbers of deaths for March 2020 in the U.S. Based on the relevant data obtained, there is an indication that the daily infection entropy can be applied in predicting the likelihood of infection at a given period. In addition, there is a relationship between the daily infection density and the time of infection. The study therefore hypothesize that, with the emergence of the new variants, average time of infection in close proximity is < 7 minutes.

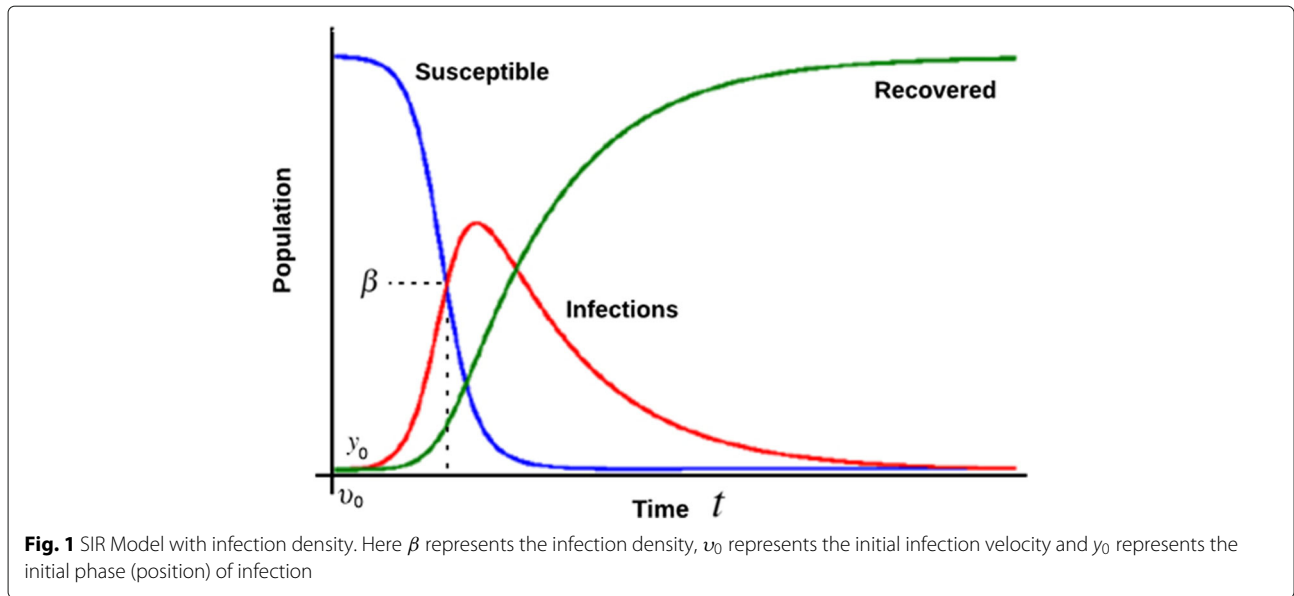
Finally, using the initial COVID-19 dataset in the U.S., this study shows how SARS-CoV-2 infection entropy can be applied in predicting the possible number of infections and deaths within a given population. Such an approach can be applicable to other disease outbreaks.

Materials and methods

In this section, we present the measures applied to evaluate the current study. First, infection density (β) is defined as the ratio of the number of infections at constant population. In this study, the unit of measurement for β is number per population. Infection acceleration is defined as the change in daily infection velocity (v_i) over time. Note, the infection acceleration, gain in virus momentum, rate of infection, and increase in the number of infections represent the same measure. These measures indicate the change in behavior of the original SARS-COV-2 virus strain that may result in transmissible new variants. Other metrics include: entropy applied to determine the uncertainties in infections and deaths. The entropy is measured in *bits* [36].

In Fig. 1, susceptible S , refers to the population who may be vulnerable to infection, infections (infected) I , are those who are infected by COVID-19, the recovered R is referred to those who with no symptom as a result of vaccines, antibodies or immune as well as those who may have died as a result of COVID-19. Considering the Susceptible Infections and Recovered (SIR) Model, we make the following assumptions:

1. A constant population with an increasing the number of infections.
2. Rate of infection in terms of infection acceleration influences the number of infections within a given population.



3. Increase in the number of infections is due to the rate of daily spread β .
4. The rate of daily spread and the infection density over a period depends on the rate at which the susceptible population is exposed to the virus and, consequently, gets infected. Hence, the rate of daily spread is equivalent to daily infection density.
5. A given population can easily be infected at close proximity with an infected individual.

The variables within the SIR model can be represented mathematically [37–39] as follows:

For the susceptible we have:

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

where

S = the susceptible,

I = infected,

α = daily reproduction rate and t = time

Assuming a decreasing number of susceptible at constant population, as S transits to I due to the rate of infections, the value of S decreases over time. Hence the value of α will remain (-ve) which shows a decrease in the number of susceptible [37–39].

For the infected we have:

$$\frac{dI}{dt} = \alpha SI - \beta I \tag{2}$$

where

S = the susceptible,

I = infected,

α = daily reproduction rate,

β = the rate of daily spread \equiv daily infection density and t = time

Assuming an increase in the number of infections due to high contact rate. Hence the value of αSI remains (+ve).

However, if the ratio of daily reproduction rate and the rate of daily spread is greater than 1, (i.e., $\frac{\alpha}{\beta} > 1$), there is every possibility that the infection will rapidly spread. On the other hand, if $\frac{\alpha}{\beta} < 1$, there could be spread with no exponential growth [37–39].

For the recovered we have:

$$\frac{dR}{dt} = \beta I \tag{3}$$

This indicates that those who have recovered from the infection due to antibodies or vaccines and may not be reinfected or even those who died. These values are excluded from the infection over time. Hence the $-\beta I$ in $\frac{dI}{dt}$ is regained as $+\beta I$ as presented in Equation 3.

where

I = infected,

R = recovered,

β = the rate of daily spread \equiv daily infection density and t = time

Computation of α , β and R_0

Note, the daily reproduction rate is computed using the expression

$$\alpha = \frac{\text{No. of infections} - \text{No. of deaths}}{\text{population size}} \times 100$$

$\frac{\alpha}{\beta} = R_0$, which represents the basic reproduction number or the basic reproductive ratio, can be assumed to

be the expected number of cases resulting from a single infection within a population where all individuals are susceptible. A higher value of R_0 means that the infection would be easily transmitted. $R_0 < 1$ means that the new cases will decrease over time, and ultimately, the outbreak will end on its own. $R_0 = 1$ means the cases may be stable over time, whereas $R_0 > 1$ indicates the virus may be autonomous, mutate into new variants, rapidly spread, and requires stringent and efficient control measures.

Looking back in March 2020, the virus's initial behavior shows how the infections will be over time with a high value of R_0 . Indicating a significant increase in positivity rate and new infection clusters, consequently increasing hospitalizations and deaths.

This implies that the daily infection density characterizes the daily reproduction rate of the SARS-CoV-2 virus. Hence,

$$f(\text{number of infections}) \Rightarrow \beta$$

The β is influenced by the number of infections within the population, which typically depends on the rate of infections. The rate of infection in this context represents the acceleration of infection. At a constant population, an increase in infection rate will lead to an increase in infection density. Consequently, if the rate (acceleration of infection) decreases, it will lead to a decrease in infection density and a decrease in the rate of daily spread.

Thus,

$$f(\text{infection acceleration}) \Rightarrow \text{number of infections}$$

The above expression implies that the increase in the number of infections is a function of infection acceleration; hence, increasing the β .

Therefore,

$$f(\text{infection acceleration}) \Rightarrow \beta$$

The infection acceleration is defined as the change in infection velocity over the change in time as follows:

$$\text{infection acceleration} = \frac{\text{change in infection velocity } (v_i)}{\text{time}} \tag{4}$$

At a constant population, the amount of infections and β within a population change as a result of infection acceleration. Thus,

$$\beta = \frac{\text{number of infections}}{\text{const. population}} \equiv \frac{f(\text{infection acceleration})}{\text{const. population}} \tag{5}$$

The β , which depends on the infection acceleration at a contact population, can be expressed as shown in Eqs. (6) and (7):

$$\beta = \frac{f(\text{infection acceleration})}{\text{const. population}} \tag{6}$$

$$\beta = \frac{\text{change in infection velocity } (v_i)}{\text{time}} \tag{7}$$

$$v_i = \beta \times \text{time} \tag{8}$$

where β = daily infection density and v_i = change in infection velocity.

By integrating the v_i over time, we obtain

$$\int v_i dt = \int \beta dt \tag{9}$$

$$\beta \int dt = \beta t + c \tag{10}$$

The initial v_i before the outbreak equal to zero (i.e., $v_i = 0$). Therefore, replacing c with this initial infection velocity, we have

$$\beta \int dt = \beta t + (v_i = 0) \tag{11}$$

To determine the stage of infection in terms of its position (y_i) within the susceptible population over time, we apply the velocity relationship [40], expressed as

$$v = \frac{\text{change in position}(y_i)}{\text{time}(t)} \tag{12}$$

where v = infection velocity.

Hence, to obtain the target infection stage y_i with respect to t , we integrate accordingly y_i with respect to t , we have

$$y_i(t) = \int v dt \tag{13}$$

Under the assumption that ($v = v_i$), by integrating the infection velocity over time, similar to Equation (9), yields

$$\int v dt = \int \beta t + v_0 dt \tag{14}$$

As we hope to determine the infection stage in terms of the position within a population, there is also a need to estimate the time these infections occur. From Equation (14) above, going further, we can determine t of infection as

$$\int \beta t + v_0 dt = \int \beta t dt + \int v_0 dt \tag{15}$$

$$= \beta \int t dt + \int v_0 dt \tag{16}$$

$$= \beta \left(\frac{1}{2} t^2 \right) + v_0 t + c \tag{17}$$

Here, c which represents the initial infection stage before the outbreak is equal to zero (i.e., $c = y_i = 0$). In

addition, at this early stage, the amount of infections as well as the infection velocity are both equal to zero.

$$y_i(t) = \beta \left(\frac{1}{2} t^2 \right) + v_0 t + (y_i = 0) \tag{18}$$

Hence, if $v = v_i$, we can now rewrite the above Equation as;

$$y_i(t) = \beta \left(\frac{1}{2} t^2 \right) + (v_i = 0)t + (y_i = 0) \tag{19}$$

$$y_i(t) = \beta \left(\frac{1}{2} t^2 \right) \tag{20}$$

To estimate the infection time (i.e., the average time it take for an infected individual to transmit the virus daily to a susceptible population at proximity), from the resulting Equation (19), we have

$$t = \sqrt{\frac{2y_i}{\beta}} \tag{21}$$

where y_i represents the stage of daily infection within the population, t represents infection time and β is the daily infection density.

Entropy

As applied in the current study, entropy can be referred to as the quantity of information uncertainties acquired from the information source measured in *bits* [36]. As the amount of uncertainties surrounding COVID-19 infection varies, the concept of entropy is applied to assess the initial daily infection uncertainties.

Thus, the entropy of an information source s on the daily infection growth rate (IGR) can be denoted by $IGR(s)$. To determine $IGR(s)$, we apply

$$IGR(s) = \sum_{i=1}^n p_i \log_2 \left(\frac{1}{p_i} \right) \tag{22}$$

We can rewrite the above equation as

$$IGR(s) = p_1 \log_2 \left(\frac{1}{p_1} \right) + \dots p_n \log_2 \left(\frac{1}{p_n} \right) \tag{23}$$

where p_i represents the probability outcome for daily infection with respect to the uncertainties surrounding the information source, $IGR(s)$ is the daily infection growth rate and n is the number of sources of information.

Justification for choice of model

Despite remarkable achievements in the fight against the virus, there are still unknown factors surrounding the COVID-19 pandemic. However, the simplicity of our proposed model reveals how important it is to consider all possible parameters that might be responsible for the increase in the number of COVID-19 cases, consequently, the new variants. In addition, it shows how the daily infection density can be modeled via the SIR model as well

as an easy to replicate approach. Notably, the current study does not involve any human or animal subjects. This study relied on the COVID-19 data reported by John Hopkins University [41], the World Health Organization [42] and the Global Initiative on Sharing All Influenza Data [43]. These datasets did not indicate the number of hospitalized persons or quarantined individuals but rather a generalized number of cases and deaths, respectively.

Data collection

The datasets applied in the current study are presented in Table 1. This study utilizes the initial COVID-19 records as reported by John Hopkins University [41], WHO [42] and GISAID [43]. Table 1 presents the number of initially reported cases with respect to the original strain of SARS-CoV-2 virus in the U.S.

Accumulated number of infections I_{cu}

The accumulated number of infections represents the possible number of COVID-19 cases over a given period (weeks or months) with respect to the average infection entropy. The accumulated number of infections is calculated as follows:

$$I_{cu(lowerlimit)} = N_{ti}(m * \left(\sum_{i=1}^n p_i \log_2 \left(\frac{1}{p_i} \right) \right)) \tag{24}$$

$$I_{cu(upperlimit)} = N_{ti}(m * \left(\sum_{i=1}^n p_i \log_2 \left(\frac{1}{p_i} \right) \right)) * (2) \tag{25}$$

where

I_{cu} = the accumulated number of infections over a given period,

N_{ti} = the total number of infected cases at a given period, m = number of months and p_i represents the probability outcome for daily infection with respect to the uncertainties surrounding the information source.

For example, if the total number of infections in the U.S., as of March 31, 2020, is 188,530 cases. To estimate the lower limit of possible number of infections in late March 2022(24 months apart), using Equation (24), we have:

$$I_{cu(lower limit)} = 188530(24(9.11)) = 41, 220, 199$$

To estimate the upper limit of infections using Equation (25), we have:

$$I_{cu(upper limit)} = 188530(24(9.11)) * 2 = 82, 440, 398$$

This number means that in late March 2022, the possible number of infections in the U.S., may be within the range of 41,220,199 and 82,440,398 but can be reduced if the necessary preventive guidelines are followed. Thus, the difference between the upper limit of infection and the possible number of cases prior to vaccine roll-out (i.e., 82,440,398 - 34,350,166 = 48,090,232). This means that

Table 1 Numbers of reported COVID-19 cases and the corresponding numbers of deaths for March 2020 in the U.S. Here *IGR* represents the infection growth rate, *DGR* represents death growth rate, and β represents the infection density measured in (num/reported cases)

March	No. of Reported Cases	Difference in the No. of Cases	IGR	No. of Deaths	Difference in the No. of Deaths	DGR	Entropy of IGR & DGR (bits)	Infection Density (β)	R_0
1	75	7	0.1	1	0	0	0.12	2.28e-7	0.985
2	100	25	0.33	6	5	5	-11.2	3.04e-7	0.936
3	124	24	0.24	9	3	0.5	0.79	3.76e-7	0.928
4	158	34	0.27	11	2	0.22	0.65	4.80e-7	0.929
5	221	63	0.4	12	1	0.09	0.59	6.71e-7	0.945
6	319	98	0.44	15	3	0.25	0.75	9.68e-7	0.953
7	435	116	0.36	19	4	0.27	0.78	1.32e-6	0.957
8	541	106	0.24	22	3	0.16	0.51	1.64e-6	0.961
9	704	163	0.3	26	4	0.18	0.55	2.14e-6	0.962
10	994	290	0.41	30	4	0.15	0.67	3.02e-6	0.969
11	1,301	307	0.31	38	8	0.27	0.68	3.95e-6	0.971
12	1,630	329	0.25	41	3	0.08	0.41	4.95e-6	0.974
13	2,183	553	0.34	48	7	0.17	0.61	6.63e-6	0.977
14	2,770	587	0.27	57	9	0.19	0.56	8.41e-6	0.979
15	3,613	843	0.3	69	12	0.21	0.64	1.09e-5	0.987
16	4,596	983	0.27	87	18	0.26	0.71	1.39e-5	0.985
17	6,344	1,748	0.38	110	23	0.26	0.78	1.93e-5	0.980
18	9,197	2,853	0.45	150	40	0.36	0.90	2.79e-5	0.984
19	13,779	4,582	0.5	206	56	0.37	0.93	4.18e-5	0.986
20	19,367	5,588	0.41	255	49	0.24	0.74	5.88e-5	0.987
21	24,192	4,825	0.25	301	46	0.18	0.50	7.34e-5	0.988
22	33,592	9,400	0.39	414	113	0.38	0.83	1.02e-4	0.987
23	43,781	10,189	0.3	555	141	0.34	0.73	1.33e-4	0.986
24	54,856	11,075	0.25	780	225	0.41	0.73	1.67e-4	0.983
25	68,211	13,355	0.24	1027	247	0.32	0.69	2.07e-4	0.985
26	85,435	17,224	0.25	1295	268	0.26	0.64	2.59e-4	0.986
27	104,126	18,691	0.22	1695	400	0.31	0.63	3.16e-4	0.984
28	123,578	19,452	0.19	2220	525	0.31	0.54	3.75e-4	0.982
29	143,491	19,913	0.16	2583	363	0.16	0.44	4.36e-4	0.981
30	163,788	20,297	0.14	3141	558	0.26	0.59	4.97e-4	0.981
31	188,530 (N_{it})	24,742	0.15	4053 (N_{id})	912	0.29	0.56	5.72e-4	0.978
			$\sum = 9.11$			$\sum = 12.45$	$\sum = 8.05$		

nearly 48 Million Americans have been prevented from COVID-19 infections and hospitalization since the vaccine rolled out. Note (34,350,166 is the upper limit of the predicted number of infections prior to the vaccines roll out in February 2021, 10 months apart).

COVID-19 average death growth rate

The average death growth rate represents the uncertainties in the number of deaths over a specified period. This metric allows us to keep track of the rate of deaths as a

result of COVID-19 over time. The average death growth rate will also help us estimate the possible number of future COVID-19 death in the U.S.

Accumulated number of deaths D_{cu}

The accumulated number of deaths represents the possible number of COVID-19 deaths cases over a given period (weeks or months) with respect to the average death entropy. The accumulated number of deaths is calculated as follows:

$$D_{cu} = N_{td}(m * \left(\sum_{i=1}^n p_i \log_2 \left(\frac{1}{p_i} \right) \right)) \quad (26)$$

where

D_{cu} = accumulated number of deaths at given period,

N_{td} = the total number of deaths at a given period,

m = number of months and p_i represents the probability outcome for daily infection with respect to the uncertainties surrounding the information source.

For example, if the number of deaths in the U.S., as of March 31, 2020, is 4053 deaths. To predict the possible number of deaths in late March 2022, using Eq. (26), we have:

$$D_{cu} = 4053(24(12.45)) = 1,211,036$$

This means with no vaccines available, in late March 2022, at least 1,211,036 deaths may be reported in the U.S., alone. However, with the available vaccines, nearly 706,437 deaths have been prevented.

Note, (504, 599 is the predicted number of deaths prior to the vaccines roll out in February 2021).

Results

The current study shows that gain in momentum of COVID-19 is influenced by the number of infections within a given population, consequently resulting in a higher daily R_0 . At constant population, again in the momentum of infection will result in a gain in infection density. Therefore, if the growth in momentum decreases, it will result in a lower infection density and a decrease in the ratio of daily reproduction rate and the rate of daily spread, respectively.

Notably, the exponential increase in the number of infections on a daily basis in March 2020, is characterized by high ratio of daily reproduction rate and the rate of daily spread $R_0 \geq 0.9$ and $R_0 < 1$ for the original SAR-COV-2 virus strain. However,

On the average infection and death growth rates, IGR achieved an entropy of 9.11 bits, whereas DGR achieved an entropy of 12.45 bits, as presented in Table 1. These uncertainties in terms of information entropy are the determinants for future forecasts on the possible number of infections and deaths. Thus, assuming no vaccines were available, in the U.S., the number of infections could range between 41,220,199 and 82,440,398, in late March 2022, with approximately, 1,211,036 deaths. However, with vaccine roll-out, approximately 48 million COVID-19 cases and 706,437 deaths have been prevented. Furthermore, the current study shows that it takes approximately 7 minutes on average for an infected individual to infect others within a susceptible population in close proximity. Hence, from the initial characteristics of the SAR-CoV-2 virus, a single person with COVID-19

can infect approximately 9 people within 1 hour and 216 people in a single day in the U.S.

Discussion

In this study, we demonstrated how the daily reproduction number of SARS-CoV-2 virus can be determined through the infection density within a given susceptible population. In addition, the current study also shows how the information entropy obtained during the early phase of the outbreak in the U.S., can be applied as determinants for predicting the number of infections and deaths. While numerous underlying but unknown factors surrounding the spread of COVID-19 still exist [44], these unknown factors may also avert the reliability of existing models in predicting and monitoring COVID-19 [45–48]. As a result, the SARS-CoV-2 virus continues to gain momentum with high stakes on human lives. Hence, it will be necessary to formulate models that can access the gain in momentum of the SARS-CoV-2 virus, which enables its ability to spread, resulting in multiple new variants [49–52].

As the need for early detection of COVID-19 infections arises, certain predictive models can be helpful in identifying potential cases [53]. For instance, a logistic model was used to predict the total number of infections to be 4 million during the outbreak in the U.S., [54]. Some of the existing models include but are not limited to the susceptible-infectious-susceptible (SIS) model, the susceptible-infected-recovered-deceased (SIRD) model alongside the SIR model, the infectious disease dynamics model and the time-dependent dynamic model previously applied in predicting the outcome of COVID-19 [44, 55–58]. An infectious disease dynamic model (SEIR) model was applied to model and predict the number of COVID-19 cases in Wuhan, China [57]. The results of that study indicate unstable values of daily reproduction rates, which may lead to a continuous increase in cases in Wuhan if public health intervention is not implemented.

A time-dependent dynamic model has been applied as a measure of public health intervention enhancement strategy through self-isolation [58]. The results obtained using the time-dependent dynamic model indicated that the daily reproduction rate of coronavirus has fallen below 1. However, the virus will continue to spread within the susceptible population [58]. On the possible time to transmit the virus from an infected person to a susceptible population, a study further reported that it might take approximately 10 mins for an infected person to produce 6,000 particles of aerosol [51]. These particles may be potentially harmful and may not be seen easily with the human eye [59].

A Markovian stochastic framework has been proposed to analyze both the reproductive ratio and the entropy

of COVID-19. These results indicated a significant but steady difference in the COVID-19 reproduction ratio and entropy, respectively, with a clear indication of the uncertainties surrounding the pandemic [60, 61]. It is therefore important to note that information entropy can be useful in differentiating between severe and mild COVID-19 patients [62–64].

However, a variation in daily reproduction ratio may vary from one location to another based on certain parameters, such as the stage of the outbreak (i.e., the rate of infection) [28]. Therefore, it is important to determine the value of the daily reproduction number at every stage of infection [28]. One of the factors responsible for the rapid transmission of COVID-19 is the daily reproduction number. However, it may be challenging to accurately determine the daily reproduction number [65, 66].

A compartmental mathematical model was formulated to predict the evolution of the virus in Cameroon and to analyze the reported cases in Brazil [67, 68]. The results achieved using these compartmental mathematical models indicated that the dynamics of COVID-19 disease are influenced by variations in the value of R_0 .

During the early phase of the pandemic, the basic reproduction ratio was estimated to range between 4.02 to 1.51 and 4.22 ± 1.69 for the U.S., and some parts of Europe, respectively, indicating a variation in the daily reproduction ratio [69, 70]. This variation in R_0 is characterized by the uncertainties surrounding the COVID-19 pandemic [26]. Hence, there is an indication that the R_0 for SARS-CoV-2 is higher than the R_0 for both SARS and MERS. As indicated earlier, previous study reported a range between 1.7 and 1.9 as the value of R_0 for SARS and $R_0 < 1$ for MERS while the R_0 for SARS-CoV-2 ranges between 2.0 and 2.5 [26].

This study also recognized some other approaches applied in uncovering useful information regarding the pandemic. For example, a study aims to properly identify critical information in an unprecedented situation such as this outbreak via a natural language processing approach to classify COVID-19-related information [71]. Such an approach enabled the extraction of certain predictor variables that can be used in predicting the amount of reposted information regarding COVID-19 on social media [71]. Similarly, a simple model constructed from the rate of social media posts can be used as a reliable prediction model when analyzing the uncertainties surrounding the pandemic [72]. Hence, accurate prediction models can be useful tools to model the outbreak of the pandemic as well as in diagnosis prediction [73, 74]. As the fight against the SAR-CoV-2 virus continues, more studies are needed to uncover useful information hidden as a result of the uncertainties surrounding the pandemic.

Despite noteworthy achievements of the existing models, there is a need to indicate how SARS-CoV-2 infection

entropy can be applied in predicting the possible number of infections and deaths. In addition, how infection density within a given population contributes to an increase in the number of cases as well as the average time it may take for an infected individual to infect a susceptible population in close proximity. Therefore, the current study is carried out to fill the gap identified above.

Conclusion and future work

The available COVID-19 vaccines have saved so many lives in the U.S., and beyond. However, several health concerns and uncertainties that have arisen in the wake of the COVID-19 pandemic have yet to be fully resolved. This study presented certain estimation models to determine the possible number of COVID-19 cases and deaths before and after vaccine roll-out in the U.S. The proposed approach shows that a high daily reproduction number for SARS-CoV-2 virus is characterized by an increase in the infection density of the original variant.

This study also shows that COVID-19 infection density can be derived via the Susceptible Infection and Recovered (SIR) model and may be applicable to other infectious diseases such as HIV. The initial behaviour of the SARS-CoV-2 virus indicates a high R_0 . Such information can be useful in monitoring the behaviour of the virus within a given period as well as in predicting possible future variants. On the projections of the pandemic in late March 2022, using the initial SAR-CoV-2 information entropy of both infection and death growth rates as the determinants for future forecasts. Assuming no vaccines available in the U.S., the current study projects that the number of infections could range between 41,220,199 and 82,440,398, with approximately 1,211,036 deaths.

Furthermore, the current study shows that it takes approximately 7 minutes on average for an infected individual to infect others within a susceptible population in close proximity. Hence, from the initial characteristics of the SAR-CoV-2 virus, this study reports that a single person with COVID-19 can infect approximately 9 people within 1 hour. Consequently, infecting about 216 people in a single day.

The proposed approach can enable other researchers to investigate other unknown factors responsible for the rapid spread of COVID-19, resulting in the emergence of new variants. This study therefore, hypothesize that the new variants (Delta and Omicron) may have a higher daily reproduction number with less daily infection entropy, consequently, spread faster and contagious.

Abbreviations

COVID-19: 2019 novel coronavirus; SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2; WHO: World Health Organization; GISAID: Global Initiative on Sharing All Influenza Data; U.S.: United States; SARS-CoV: Severe acute respiratory syndrome coronavirus; MERS-CoV: Middle East respiratory syndrome coronavirus; GDP: Gross domestic product; SIR: Susceptible Infections and Recovered; IGR: Infection growth rate; SIS: Susceptible-

infectious-susceptible; SIRD: Susceptible- infected-recovered-deceased; SARS: Severe Acute Respiratory Syndrome; MERS: Middle East Respiratory Syndrome; HIV: Human immunodeficiency virus

Acknowledgements

The authors would like to thank the anonymous reviewers for their time and thoughtful comments while reviewing our manuscript.

Authors' contributions

Ebubeogu Amarachukwu Felix and Paulinus Ofem design the research and the proposed method, Chamberline Ekene Ozigbu, Azizi Seixas and Khoulood Maswadi performed data curation, data preprocessing and formal analysis, respectively. Donaldson F. Conserve secured the funding, supervision and corrections. Ebubeogu Amarachukwu Felix and Paulinus Ofem drafted the manuscript, and all co-authors reviewed the manuscript. The authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Software Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia. ²Department of Health Services Policy and Management, Arnold School of Public Health, 29208 Columbia, SC, United States. ³Department of Prevention and Community Health, Milken Institute School of Public Health, The George Washington University, 20052 Washington, United States. ⁴Department of Psychiatry and Behavioral Sciences, The University of Miami Miller School of Medicine, 33136 Miami, FL, United States. ⁵Department of Software Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia. ⁶Department of Management Information Systems, Jazan University, 45142 Jazan, Saudi Arabia.

Received: 3 January 2022 Accepted: 3 March 2022

Published online: 28 March 2022

References

1. Adalja AA, Toner E, Inglesby TV. Priorities for the us health community responding to covid-19. *Jama*. 2020;323(14):1343–4.
2. Wu D, Wu T, Liu Q, Yang Z. The sars-cov-2 outbreak: what we know. *Int J Infect Dis*. 2020;94:44–8.
3. Hu C, Liu Z, Jiang Y, Shi O, Zhang X, Xu K, Suo C, Wang Q, Song Y, Yu K. Early prediction of mortality risk among patients with severe COVID-19, using machine learning. *Intl J Epidemiol*. 2020;49(6):1918–29.
4. Gee J, Marquez P, Su J, Calvert GM, Liu R, Myers T, Nair N, Martin S, Clark T, Markowitz L, et al. First month of covid-19 vaccine safety monitoring—united states, december 14, 2020–january 13, 2021. *Morb Mortal Wkly Rep*. 2021;70(8):283.
5. Ye Q, Zhou J, Wu H, et al. Using information technology to manage the covid-19 pandemic: development of a technical framework based on practical experience in china. *JMIR Med Inf*. 2020;8(6):19515.
6. Kumar A, Gupta PK, Srivastava A. A review of modern technologies for tackling covid-19 pandemic. *Diabetes Metab Syndr Clin Res Rev*. 2020;14(4):569–73.
7. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and covid-19. *Nat Med*. 2020;26(4):459–61.
8. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, Fan Y, Zheng C. Radiological findings from 81 patients with covid-19 pneumonia in wuhan, china: a descriptive study. *Lancet Infect Dis*. 2020;20(4):425–34.
9. Wei X-S, Wang X-R, Zhang J-C, Yang W-B, Ma W-L, Yang B-H, Jiang N-C, Gao Z-C, Shi H-Z, Zhou Q. A cluster of health care workers with covid-19 pneumonia caused by sars-cov-2. *J Microbiol Immunol Infect*. 2021;54(1):54–60.
10. Alazawy A, Arshad S-S, Bejo M-H, Omar A-R, Tengku Ibrahim T-A, Sharif S, Bande F, Awang-Isa K. Ultrastructure of felis catus whole fetus (fcwf-4) cell culture following infection with feline coronavirus. *J Electron Microsc*. 2011;60(4):275–82.
11. Hwa K-Y, Lin WM, Hou Y-I, Yeh T-M. Molecular mimicry between sars coronavirus spike protein and human protein. In: 2007 *Frontiers in the Convergence of Bioscience and Information Technologies*. IEEE; 2007. p. 294–8.
12. Peiris J, Lai S, Poon L, Guan Y, Yam L, Lim W, Nicholls J, Yee W, Yan W, Cheung M, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*. 2003;361(9366):1319–25.
13. Xiong X, Chua GT, Chi S, Kwan MYW, Wong WHS, Zhou A, Shek CC, Tung KT, Qin H, Wong RS, et al. A comparison between chinese children infected with coronavirus disease-2019 and with severe acute respiratory syndrome 2003. *J Pediatr*. 2020;224:30–6.
14. Oxford J, Bossuyt S, Lambkin R. A new infectious disease challenge: Urbani severe acute respiratory syndrome (sars) associated coronavirus. *Immunology*. 2003;109(3):326.
15. Li K, Wohlford-Lenane C, Perlman S, Zhao J, Jewell AK, Reznikov LR, Gibson-Corley KN, Meyerholz DK, McCray Jr PB. Middle east respiratory syndrome coronavirus causes multiple organ damage and lethal disease in mice transgenic for human dipeptidyl peptidase 4. *J Infect Dis*. 2016;213(5):712–22.
16. Chan JF-W, Lau SK-P, Woo PC-Y. The emerging novel middle east respiratory syndrome coronavirus: the “knowns” and “unknowns”. *J Formos Med Assoc*. 2013;112(7):372–81.
17. Elkholy AA, Grant R, Assiri A, Elhakim M, Malik MR, Van Kerkhove MD. Mers-cov infection among healthcare workers and risk factors for death: retrospective analysis of all laboratory-confirmed cases reported to who from 2012 to 2 june 2018. *J Infect Public Health*. 2020;13(3):418–22.
18. Ullah A, Mabood N, Maqbool M, Khan L, Khan M, Ullah M. Sar-cov-2 infection, emerging new variants and the role of activation induced cytidine deaminase (aid) in lasting immunity. *Saudi Pharm J*. 2021;29(10):1181–4.
19. Mahase E. Covid-19: Hospital admission 50–70% less likely with omicron than delta, but transmission a major concern: *British Medical Journal Publishing Group*; 2021.
20. Farinholt T, Doddapaneni H, Qin X, Menon V, Meng Q, Metcalf G, Chao H, Gingras M-C, Avadhanula V, Farinholt P, et al. Transmission event of sars-cov-2 delta variant reveals multiple vaccine breakthrough infections. *BMC Med*. 2021;19(1):1–6.
21. Mohapatra RK, Sarangi AK, Kandi V, Azam M, Tiwari R, Dhama K. Omicron (b. 1.1. 529 variant of sars-cov-2); an emerging threat: current global scenario. *J Med Virol*. 2021;2022:1–4.
22. Rabaan AA, Al-Ahmed SH, Haque S, Sah R, Tiwari R, Malik YS, Dhama K, Yattoo MI, Bonilla-Aldana DK, Rodriguez-Morales AJ, et al. Sars-cov-2, sars-cov, and mers-cov: a comparative overview. *Infez Med*. 2020;28(2):174–84.
23. Hassanzadeh K, Perez Pena H, Dragotto J, Buccarello L, Iorio F, Pieraccini S, Sancini G, Feligioni M. Considerations around the sars-cov-2 spike protein with particular attention to covid-19 brain infection and neurological symptoms. *ACS Chem Neurosci*. 2020;11(15):2361–9.
24. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, et al. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New Engl J Med*. 2020;382:1–9.
25. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *Lancet*. 2020;395(10225):689–97.
26. Petrosillo N, Viceconte G, Ergonul O, Ippolito G, Petersen E. Covid-19, sars and mers: are they closely related?. *Clin Microbiol Infect*. 2020;26(6):729–34.
27. Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, Gao D, Yang L, He D, et al. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-ncov) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int J Infect Dis*. 2020;92:214–7.

28. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of covid-19 is higher compared to sars coronavirus. *J Travel Med.* 2020;27:1–4.
29. Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP. Novel coronavirus 2019-nCoV (COVID-19): early estimation of epidemiological parameters and epidemic size estimates. *Philosophical Trans R Soc B.* 2021;376(1829):20200265.
30. Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP. Novel coronavirus 2019-ncov (covid-19): early estimation of epidemiological parameters and epidemic size estimates. *Phil Trans R Soc B.* 2021;376(1829):20200265.
31. Organization WH, et al. Consensus document on the epidemiology of severe acute respiratory syndrome (sars). Technical report, World Health Organization. 2003.
32. Ke R, Romero-Severson E, Sanche S, Hengartner N. Estimating the reproductive number r_0 of sars-cov-2 in the united states and eight european countries and implications for vaccination. *J Theor Biol.* 2021;517:110621.
33. Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, Wu J. Estimation of the transmission risk of the 2019-ncov and its implication for public health interventions. *J Clin Med.* 2020;9(2):462.
34. Li Y-D, Chi W-Y, Su J-H, Ferrall L, Hung C-F, Wu T-C. Coronavirus vaccine development: from sars and mers to covid-19. *J Biomed Sci.* 2020;27(1):1–23.
35. Chakraborty I, Maity P. COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Sci Total Environ.* 2020;728:138882.
36. Felix EA, Lee SP. Predicting the number of defects in a new software version. *PLoS ONE.* 2020;15(3):0229131.
37. Dallas TA, Carlson CJ, Poisot T. Testing predictability of disease outbreaks with a simple model of pathogen biogeography. *R Soc Open Sci.* 2019;6(11):190883.
38. De Groot M, Ogris N. Short-term forecasting of bark beetle outbreaks on two economically important conifer tree species. *For Ecol Manag.* 2019;450:117495.
39. Kelly JD, Park J, Harrigan RJ, Hoff NA, Lee SD, Wannier R, Selo B, Mossoko M, Njoloko B, Okitolonda-Wemakoy E, et al. Real-time predictions of the 2018–2019 ebola virus disease outbreak in the democratic republic of the congo using hawkes point process models. *Epidemics.* 2019;28:100354.
40. Felix EA, Lee SP. Integrated approach to software defect prediction. *IEEE Access.* 2017;5:21524–47.
41. Johns Hopkins University & Medicine. Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2020. <https://www.coronavirus.jhu.edu/map.html>. Accessed 02 Apr 2020.
42. World Health Organization. Coronavirus Disease (COVID-19) Outbreak Situation. 2020. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Accessed 02 Apr 2020.
43. Gostin LO. hCoV-19 Tracking of Variants. 2020. <https://www.gisaid.org/hcov19-variants>. Accessed 02 July 2020.
44. Pinter G, Felde I, Mosavi A, Ghamisi P, Gloaguen R. Covid-19 pandemic prediction for hungary; a hybrid machine learning approach. *Mathematics.* 2020;8(6):890.
45. Rypdal M, Sugihara G. Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics. *Nat Commun.* 2019;10(1):1–8.
46. Scarpino SV, Petri G. On the predictability of infectious disease outbreaks. *Nat Commun.* 2019;10(1):1–8.
47. Zhan Z, Dong W, Lu Y, Yang P, Wang Q, Jia P. Real-time forecasting of hand-foot-and-mouth disease outbreaks using the integrating compartment model and assimilation filtering. *Sci Rep.* 2019;9(1):1–9.
48. Luo J. Predictive monitoring of covid-19. *SUTD Data-Driven Innov Lab.* 2020;446:1–12.
49. Kannan SR, Spratt AN, Cohen AR, Naqvi SH, Chand HS, Quinn TP, Lorson CL, Byrareddy SN, Singh K. Evolutionary analysis of the delta and delta plus variants of the sars-cov-2 viruses. *J Autoimmun.* 2021;124:102715.
50. Del Rio C, Malani PN, Omer SB. Confronting the delta variant of sars-cov-2, summer 2021. *Jama.* 2021;326(11):1001–1002.
51. Asadi S, Bouvier N, Wexler AS, Ristenpart WD. The coronavirus pandemic and aerosols: Does COVID-19 transmit via expiratory particles? *Taylor & Francis*; 2020.
52. Kupferschmidt K, Wadman M. Delta variant triggers new phase in the pandemic: American Association for the Advancement of Science; 2021.
53. Poletto C, Scarpino SV, Volz EM. Applications of predictive modelling early in the covid-19 epidemic. *Lancet Dig Health.* 2020;2(10):498–9.
54. Bhardwaj R. A predictive model for the evolution of covid-19. *Trans Indian Natl Acad Eng.* 2020;5(2):133–40.
55. Miller JC. A note on the derivation of epidemic final sizes. *Bull Math Biol.* 2012;74(9):2125–41.
56. Miller JC. Mathematical models of sir disease spread with combined non-sexual and sexual transmission routes. *Inf Dis Model.* 2017;2(1):35–55.
57. Wang H, Wang Z, Dong Y, Chang R, Xu C, Yu X, Zhang S, Tsamslag L, Shang M, Huang J, et al. Phase-adjusted estimation of the number of coronavirus disease 2019 cases in wuhan, china. *Cell Discov.* 2020;6(1):1–8.
58. Tang B, Bragazzi NL, Li Q, Tang S, Xiao Y, Wu J. An updated estimation of the risk of transmission of the novel coronavirus (2019-ncov). *Infect Dis Model.* 2020;5:248–55.
59. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD. Aerosol emission and superemission during human speech increase with voice loudness. *Sci Rep.* 2019;9(1):1–10.
60. Wang Z, Broccardo M, Mignan A, Sornette D. The dynamics of entropy in the covid-19 outbreaks. *Nonlinear Dyn.* 2020;101(3):1847–69.
61. Bandt C. Entropy ratio and entropy concentration coefficient, with application to the covid-19 pandemic. *Entropy.* 2020;22(11):1315.
62. Bajić D, Dajić V, Milovanović B. Entropy analysis of covid-19 cardiovascular signals. *Entropy.* 2021;23(1):87.
63. Albahri AS, Hamid RA, Albahri OS, Zaidan A. Detection-based prioritisation: Framework of multi-laboratory characteristics for asymptomatic covid-19 carriers based on integrated entropy–topsis methods. *Artif Intell Med.* 2021;111:101983.
64. Hasan AM, Al-Jawad MM, Jalab HA, Shaiba H, Ibrahim RW, AL-Shamasneh AR. Classification of covid-19 coronavirus, pneumonia and healthy lungs in ct scans using q-deformed entropy and deep learning features. *Entropy.* 2020;22(5):517.
65. Viceconte G, Petrosillo N. COVID-19 R0: Magic number or conundrum? *Multidisciplinary Digital Publishing Institute*; 2020.
66. Tao Y. Maximum entropy method for estimating the reproduction number: An investigation for covid-19 in china and the united states. *Phys Rev E.* 2020;102(3):032136.
67. Nabi KN, Abboubakar H, Kumar P. Forecasting of covid-19 pandemic: From integer derivatives to fractional derivatives. *Chaos Solitons Fractals.* 2020;141:110283.
68. Kumar P, Erturk VS, Abboubakar H, Nisar KS. Prediction studies of the epidemic peak of coronavirus disease in brazil via new generalised caputo type fractional derivatives. *Alex Eng J.* 2021;60(3):3189–204.
69. Gunzler D, Sehgal AR. Optimal control of fractional order COVID-19 epidemic spreading in Japan and India 2020. *Biophys Rev Letters.* 2020;15(04):207–236.
70. Linka K, Peirlinck M, Kuhl E. The reproduction number of covid-19 and its correlation with public health interventions. *Comput Mech.* 2020;66(4):1035–50.
71. Li L, Zhang Q, Wang X, Zhang J, Wang T, Gao T, Duan W, Tsoi KK, Wang F. Characterizing the propagation of situational information in social media during covid-19 epidemic: A case study on weibo. *IEEE Trans Comput Soc Syst.* 2020;7(2):556–62.
72. Asur S, Huberman BA. Predicting the future with social media. In: 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 1. IEEE; 2010. p. 492–9.
73. Ardabili SF, Mosavi A, Ghamisi P, Ferdinand F, Varkonyi-Koczy AR, Reuter U, Rabczuk T, Atkinson PM. Covid-19 outbreak prediction with machine learning. *Algorithms.* 2020;13(10):249.
74. Zoabi Y, Deri-Rozov S, Shomron N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *npj Digit Med.* 2021;4(1):1–5.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.