Research article

A mathematical model for the novel coronavirus epidemic in Wuhan, China

Chayu Yang and Jin Wang[*]

Department of Mathematics, University of Tennessee at Chattanooga, 615 McCallie Ave., Chattanooga, TN 37403, USA

* **Correspondence:** Email: Jin-Wang02@utc.edu; Tel: +14234254725.

Abstract: We propose a mathematical model to investigate the current outbreak of the coronavirus disease 2019 (COVID-19) in Wuhan, China. Our model describes the multiple transmission pathways in the infection dynamics, and emphasizes the role of the environmental reservoir in the transmission and spread of this disease. Our model also employs non-constant transmission rates which change with the epidemiological status and environmental conditions and which reflect the impact of the on-going disease control measures. We conduct a detailed analysis of this model, and demonstrate its application using publicly reported data. Among other findings, our analytical and numerical results indicate that the coronavirus infection would remain endemic, which necessitates long-term disease prevention and intervention programs.

Keywords: COVID-19; compartmental modeling; basic reproduction number

1. Introduction

A severe outbreak of respiratory illness started in Wuhan, a city of 11 million people in central China, in December 2019. The causative agent is the novel coronavirus which was identified and isolated from a single patient in early January and subsequently verified in 16 additional patients [1]. The virus is believed to have a zoonotic origin. In particular, the Huanan Seafood Market, a live animal and seafood wholesale market in Wuhan, was regarded as a primary source of this epidemic, as it is found that 55% of the first 425 confirmed cases were linked to the marketplace [2]. Meanwhile, recent comparisons of the genetic sequences of this virus and bat coronaviruses show a 96% similarity [3]. This is the third zoonotic human coronavirus emerging in the current century, after the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2002 that spread to 37 countries and the Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012 that spread to 27 countries. Typical symptoms of COVID-19 infection include dry cough, fever, fatigue, breathing difficulty, and bilateral lung infiltration in severe cases, similar to those caused by SARS-CoV and MERS-CoV infections [4].

Some people also develop non-respiratory symptoms such as nausea, vomiting, and diarrhea [5, 6].

The COVID-19 outbreak is currently on-going and the number of infections has been fast growing since the onset of the epidemic. By January 23, 571 confirmed cases had been reported in China and the infections were already exported to the US, Thailand, Japan and Republic of Korea [7,8]. As a bold effort to contain the epidemic, the Chinese government ordered to lock down Wuhan on January 23 and at least 15 other cities in the following days, effectively restricting the movement of more than 50 million people in central China, which is considered as the largest quarantine in human history. As of February 10, more than 42,600 people in China have been diagnosed with coronavirus [7]. On January 30, the World Health Organization (WHO) formally declared the outbreak of novel coronavirus a Global Public Health Emergency of International Concern [9].

Several factors complicate the infection dynamics of COVID-19 and add challenges to the disease control. First, the origin of the infection is still uncertain, although it is widely speculated that wild animals such as bats, civets and minks are responsible for starting the epidemic [3]. Second, clinical evidence shows that the incubation period of this disease ranges from 2 to 14 days. During this period of time, infected individuals may not develop any symptoms and may not be aware of their infection, yet they are capable of transmitting the disease to other people [10]. Third, the virus is new and there are no antiviral drugs or vaccines currently available. Consequently, disease control heavily relies on prompt detection and isolation of symptomatic cases. Additionally, the disease outbreak took place right before the Spring Festival (i.e., the Lunar New Year), the most important holiday in China, and a huge population (several millions for the city of Wuhan alone, and hundreds of millions on a national scale) traveled during the month of January 2020, which makes fast and wide spread of the infection possible.

A number of modeling studies have already been performed for the COVID-19 epidemic. Wu et al. [11] introduced a susceptible-exposed-infectious-recovered (SEIR) model to describe the transmission dynamics, and forecasted the national and global spread of the disease, based on reported data from December 31, 2019 to January 28, 2020. They also estimated that the basic reproductive number for COVID-19 was about 2.68. Read et al. [12] reported a value of 3.1 for the basic reproductive number based on data fitting of a SEIR model, using an assumption of Poisson-distributed daily time increments. Tang et al. [13] proposed a deterministic compartmental model incorporating the clinical progression of the disease, the individual epidemiological status, and the intervention measures. They found that the control reproductive number could be as high as 6.47, and that intervention strategies such as intensive contact tracing followed by quarantine and isolation can effectively reduce the control reproduction number and the transmission risk. Imai et al. [14] conducted computational modeling of potential epidemic trajectories to estimate the size of the disease outbreak in Wuhan, with a focus on the human-to-human transmission. Their results imply that control measures need to block well over 60% of transmission to be effective in containing the outbreak. In addition, Gao et al. [15] developed a deep learning algorithm to analyze the infectivity of the novel coronavirus and predict its potential hosts. Their results indicate that bats and minks may be two animal hosts of this virus.

Most of these models have emphasized the significant role of the direct, human-to-human transmission pathway in this epidemic [16], as highlighted by the facts that the majority of the infected individuals did not have any contact with the marketplaces in Wuhan, that the number of infections has been rapidly increasing, and that the disease has spread to all provinces in China as well as more than 20 other countries. In particular, a large number of infected individuals exhibit a relatively long

incubation period so that they do not show any symptoms and are unaware of their infection for as long as 10–14 days, during which time they can easily transmit the disease to other people through direct contact.

On the other hand, the models published thus far have not taken into account the role of the environment in the transmission of COVID-19. For example, it is reported that environmental samples taken from the areas of the Huanan Seafood Market have come back positive for the novel coronavirus [4], suggesting that the pathogen may be transmitted through the environmental reservoir. When infected individuals cough or sneeze, they may spread the virus to the environment through their respiratory droplets which then may go on to infect other people with close contact of the same area. Such transmission would especially be facilitated during the early period of the disease outbreak when the general public was not aware of the infection risk, infected individuals were not isolated, and most people did not wear face masks. Even worse, there is a possibility that the virus may survive in the environment for several days, increasing the risk of contamination via surfaces and fomites [17, 18]. Such environmental survival was confirmed for SARS-CoV [19]. A most recent study, based on the review of 22 types of coronaviruses, reveals that coronaviruses such as SARS-CoV, MERS-CoV and endemic human coronaviruses can persist on inanimate surfaces like metal, glass or plastic for up to 9 days [20], providing strong evidences for the pathogen's environmental survival. Additionally, the novel coronavirus has been found in the stool of some infected individuals [5], which may contaminate the aquatic environment, and fecal-oral contact remains a possible route of transmission for this disease.

In the present paper, we present a new mathematical model for COVID-19 that incorporates multiple transmission pathways, including both the environment-to-human and human-to-human routes. In particular, we introduce an environmental compartment that represents the pathogen concentration in the environmental reservoir. A susceptible individual may contract the disease through the interaction with the contaminated environment, with an infectious but asymptomatic individual, or with an infectious and symptomatic individual. Meanwhile, the transmission rates in our model depend on the epidemiological status and environmental conditions which change with time. In particular, when the infection level is high, people would be motivated to take necessary action to reduce the contact with the infected individuals and contaminated environment so as to protect themselves and their families, leading to a reduction of the average transmission rates. Such varied transmission rates also reflect the strong disease control measures that the Chinese government has implemented, including large-scale quarantine, intensive tracking of movement and contact, strict isolation, extending the Lunar New Year holiday, and advising the public to stay home and avoid spreading infection.

The remainder of this paper is organized as follows. In Section 2, we present our model and assumptions, and conduct a detailed mathematical analysis. In Section 3, we conduct numerical simulation by incorporating the infection data reported for the city of Wuhan. In Section 4, we conclude the paper with some discussion.

2. Model formulation and analysis

2.1. Formulation

We divide the total human population into four compartments: the susceptible (denoted by S), the exposed (denoted by E), the infected (denoted by I), and the recovered (denoted by R). Individuals in the infected class have fully developed disease symptoms and can infect other people. Individuals

in the exposed class are in the incubation period; they do not show symptoms but are still capable of infecting others. Thus, another interpretation of the E and I compartments in our model is that they contain asymptomatic infected and symptomatic infected individuals, respectively.

We introduce the following model to describe the transmission dynamics of the COVID-19 epidemic:

$$\frac{dS}{dt} = \Lambda - \beta_E(E)SE - \beta_I(I)SI - \beta_V(V)SV - \mu S,$$

$$\frac{dE}{dt} = \beta_E(E)SE + \beta_I(I)SI + \beta_V(V)SV - (\alpha + \mu)E,$$

$$\frac{dI}{dt} = \alpha E - (w + \gamma + \mu)I, \tag{2.1}$$

$$\frac{dR}{dt} = \gamma I - \mu R,$$

$$\frac{dV}{dt} = \xi_1 E + \xi_2 I - \sigma V,$$

where V is the concentration of the coronavirus in the environmental reservoir. The parameter Λ represents the population influx, μ is the natural death rate of human hosts, α^{-1} is the incubation period between the infection and the onset of symptoms, w is the disease-induced death rate, γ is the rate of recovery from infection, ξ_1 and ξ_2 are the respective rates of the exposed and infected individuals contributing the coronavirus to the environmental reservoir, and σ is the removal rate of the virus from the environment. The functions $\beta_E(E)$ and $\beta_I(I)$ represent the direct, human-to-human transmission rates between the exposed and susceptible individuals, and between the infected and susceptible individuals, respectively, and the function $\beta_V(V)$ represents the indirect, environment-to-human transmission rate. We assume that $\beta_E(E)$, $\beta_I(I)$ and $\beta_V(V)$ are all non-increasing functions, given that higher values of E, I and V would motivate stronger control measures that could reduce the transmission rates. Specifically, we make the following assumptions:

(A1) $\beta_E(E), \beta_I(I), \beta_V(V)$ are all positive; and
(A2) $\beta'_E(E) \leq 0, \beta'_I(I) \leq 0, \beta'_V(V) \leq 0$.

Apparently, system (2.1) has a unique disease-free equilibrium (DFE) at

$$X_0 = (S_0, E_0, I_0, R_0, V_0) = \left(\frac{\Lambda}{\mu}, 0, 0, 0, 0\right). \tag{2.2}$$

The infection components in this model are E, I, and V. The new infection matrix F and the transition matrix V are given by

$$F = \begin{bmatrix} \beta_E(0)S_0 & \beta_I(0)S_0 & \beta_V(0)S_0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \alpha + \mu & 0 & 0 \\ -\alpha & w_1 & 0 \\ -\xi_1 & -\xi_2 & \sigma \end{bmatrix}, \tag{2.3}$$

where $w_1 = w + \gamma + \mu$. The basic reproduction number of model (2.1) is then defined as the spectral radius of the next generation matrix FV^{-1} [21]; i.e.,

$$\mathcal{R}_0 = \rho(FV^{-1}) = \frac{\beta_E(0)S_0}{\alpha + \mu} + \frac{\alpha\beta_I(0)S_0}{w_1(\alpha + \mu)} + \frac{(w_1\xi_1 + \alpha\xi_2)\beta_V(0)S_0}{\sigma w_1(\alpha + \mu)}$$

$$:= \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3, \tag{2.4}$$

which provides a quantification of the disease risk. The first two parts $\mathcal{R}_1$ and $\mathcal{R}_2$ measure the contributions from the human-to-human transmission routes (exposed-to-susceptible and infected-to-susceptible, respectively), and the third part $\mathcal{R}_3$ represents the contribution from the environment-to-human transmission route. These three transmission modes collectively shape the overall infection risk for the COVID-19 outbreak.

2.2. Equilibrium analysis

We now analyze the equilibria of the system (2.1) which will provide essential information regarding the long-term dynamics of the disease. Let (S, E, I, R, V) be an equilibrium of model (2.1) and thereby satisfy the following equations

$$
\begin{aligned}
\Lambda - \beta_E(E)SE - \beta_I(I)SI - \beta_V(V)SV - \mu S &= 0, \\
\beta_E(E)SE + \beta_I(I)SI + \beta_V(V)SV - (\alpha + \mu)E &= 0, \\
\alpha E - w_1 I &= 0, \\
\gamma I - \mu R &= 0, \\
\xi_1 E + \xi_2 I - \sigma V &= 0.
\end{aligned}
\tag{2.5}
$$

Solving (2.5) yields

$$
\begin{aligned}
S &= \frac{1}{\mu}\left(\Lambda - (\alpha + \mu)E\right), \\
E &= \frac{w_1}{\alpha}I, \\
R &= \frac{\gamma}{\mu}I, \\
V &= \frac{w_1\xi_1 + \alpha\xi_2}{\sigma\alpha}I.
\end{aligned}
\tag{2.6}
$$

It follows from the first two equations of (2.6) that S can be denoted by a function of I, namely,

$$
S = \phi(I) := \frac{1}{\mu}\left(\Lambda - \frac{w_1(\alpha + \mu)}{\alpha}I\right).
\tag{2.7}
$$

Meanwhile, in view of the second equation of (2.5) and Eqs (2.6), we obtain

$$
S = \psi(I) := (\alpha + \mu)\left(\beta_E\left(\frac{w_1}{\alpha}I\right) + \frac{\alpha}{w_1}\beta_I(I) + \frac{w_1\xi_1 + \alpha\xi_2}{\sigma w_1}\beta_V\left(\frac{w_1\xi_1 + \alpha\xi_2}{\sigma\alpha}I\right)\right)^{-1}.
\tag{2.8}
$$

Let us now consider curves $S = \phi(I), \ I \geq 0$ and $S = \psi(I), \ I \geq 0$. In particular, the intersections of these two curves in $\mathbb{R}_+^2$ determine the non-DFE equilibria. Clearly, $\phi(I)$ is strictly decreasing, whereas $\psi(I)$ is increasing since $\beta_E\left(\frac{w_1}{\alpha}I\right)$, $\beta_I(I)$, and $\beta_V\left(\frac{w_1\xi_1 + \alpha\xi_2}{\sigma\alpha}I\right)$ are positive and decreasing functions of I. Additionally, one can easily verify that $\phi(0) = S_0$, $\phi(I_1) = 0$, where $I_1 = \frac{\alpha\Lambda}{(\alpha+\mu)w_1}$, and

$$
\psi(0) = (\alpha + \mu)\left(\beta_E(0) + \frac{\alpha}{w_1}\beta_I(0) + \frac{w_1\xi_1 + \alpha\xi_2}{\sigma w_1}\beta_V(0)\right)^{-1} = \frac{S_0}{\mathcal{R}_0}.
$$

Thus, we conclude:

(1) If $\mathcal{R}_0 > 1$, these two curves have a unique intersection lying in the interior of $\mathbb{R}_+^2$, since $\psi(0) < \phi(0)$ and $\psi(I_1) \geq \psi(0) > 0 = \phi(I_1)$. Furthermore, at this intersection point, Eq (2.6) yields a unique endemic equilibrium (EE)

$$X_* = (S_*, E_*, I_*, R_*, V_*).$$

(2) If $\mathcal{R}_0 \leq 1$, the two curves have no intersection in the interior of $\mathbb{R}_+^2$ as $\psi(0) \geq \phi(0)$.

Therefore, by Eq (2.6), we find that the model (2.1) admits a unique equilibrium, the DFE X_0, if $\mathcal{R}_0 \leq 1$; and it admits two equilibria, the DFE X_0 and the EE X_*, if $\mathcal{R}_0 > 1$.

In what follows, we perform a study on the global stability of the DFE. By a simple comparison principle, we find that $0 \leq S + E + I + R \leq S_0$ and $0 \leq V \leq \frac{(\xi_1+\xi_2)S_0}{\sigma}$. Thus, it leads to a biologically feasible domain

$$\Omega = \left\{ (S, E, I, R, V) \in \mathbb{R}_+^5 : S + E + I + R \leq S_0,\, 0 \leq V \leq \frac{(\xi_1 + \xi_2)S_0}{\sigma} \right\}.$$

Theorem 2.1. *The following statements hold for the model* (2.1).

(1) *If $\mathcal{R}_0 \leq 1$, the DFE of system* (2.1) *is globally asymptotically stable in Ω.*
(2) *If $\mathcal{R}_0 > 1$, the DFE of system* (2.1) *is unstable and there exists a unique endemic equilibrium. Moreover, the disease is uniformly persistent in the interior of Ω, denoted by $\mathring{\Omega}$; namely,* $\liminf\limits_{t \to \infty}(E(t), I(t), V(t)) > (\varepsilon, \varepsilon, \varepsilon)$ *for some $\varepsilon > 0$.*

Proof. Let $\mathbf{X} = (E, I, V)^T$. One can verify that

$$\frac{d\mathbf{X}}{dt} \leq (F - V)\mathbf{X},$$

where the matrices F and V are given in Eq (2.3). By manipulating some algebraic computaion, we let $\mathbf{u} = (\beta_E(0), \beta_I(0), \beta_V(0))$. It then follows from the fact $\mathcal{R}_0 = \rho(FV^{-1}) = \rho(V^{-1}F)$ and direct calculation that $\mathbf{u}$ is a left eigenvector associated with the eigenvalue $\mathcal{R}_0$ of the matrix $V^{-1}F$; i.e., $\mathbf{u}V^{-1}F = \mathcal{R}_0\mathbf{u}$. Consider a Lyapunov function

$$\mathcal{L}_0 = \mathbf{u}V^{-1}\mathbf{X}.$$

Differentiating $\mathcal{L}$ along the solutions of (2.1), we have

$$\frac{d\mathcal{L}_0}{dt} = \mathbf{u}V^{-1}\frac{d\mathbf{X}}{dt} \leq \mathbf{u}V^{-1}(F - V)\mathbf{X} = \mathbf{u}(\mathcal{R}_0 - 1)\mathbf{X}. \tag{2.9}$$

If $\mathcal{R}_0 < 1$, the equality $\frac{d\mathcal{L}_0}{dt} = 0$ implies that $\mathbf{u}\mathbf{X} = 0$. This leads to $E = I = V = 0$ by noting that all components of $\mathbf{u}$ are positive. Hence, when $\mathcal{R}_0 < 1$, equations of (2.5) yield $S = S_0$, and $E = I = R = V = 0$. Thus, the invariant set on which $\frac{d\mathcal{L}_0}{dt} = 0$ contains only the point X_0.

If $\mathcal{R}_0 = 1$, then the equality $\frac{d\mathcal{L}_0}{dt} = 0$ implies that

$$\left(\frac{\beta_E(E)S}{S_0} + \frac{\alpha\beta_I(I)}{w_1} + \frac{(w_1\xi_1 + \alpha\xi_2)\beta_V(V)}{\sigma w_1} - \frac{\alpha + \mu}{S_0} \right)E$$
$$+ \left(\frac{S}{S_0}\beta_I(I) - \beta_I(0) \right)I + \left(\frac{S}{S_0}\beta_V(V) - \beta_V(0) \right)V = 0.$$

It is easy to see that

$$\frac{S}{S_0}\beta_I(I) - \beta_I(0) \le 0, \quad \frac{S}{S_0}\beta_V(V) - \beta_V(0) \le 0,$$

and

$$\frac{\beta_E(E)S}{S_0} + \frac{\alpha\beta_I(I)}{w_1} + \frac{(w_1\xi_1 + \alpha\xi_2)\beta_V(V)}{\sigma w_1} - \frac{\alpha + \mu}{S_0}$$

$$\le \frac{\alpha + \mu}{S_0}\left(\frac{\beta_E(0)S_0}{\alpha + \mu} + \frac{\alpha\beta_I(0)S_0}{w_1(\alpha + \mu)} + \frac{(w_1\xi_1 + \alpha\xi_2)\beta_V(0)S_0}{\sigma w_1(\alpha + \mu)} - 1\right)$$

$$= \frac{\alpha + \mu}{S_0}(\mathcal{R}_0 - 1) = 0.$$

Hence, we have either $E = I = V = 0$, or $\beta_E(E) = \beta_E(0), \beta_I(I) = \beta_I(0), \beta_V(V) = \beta_V(0)$, and $S = S_0$. As processed before, each of cases would indicate the DEF X_0 is the only invariant set on $\{(S, E, I, R, V) \in \Omega : \frac{d\mathcal{L}_0}{dt} = 0\}$.

Therefore, when $\mathcal{R}_0 < 1$ or $\mathcal{R}_0 = 1$, the largest invariant set on which $\frac{d\mathcal{L}_0}{dt} = 0$ always consists of the singleton $X_0 = (S_0, 0, 0, 0, 0)$. By LaSalle's Invariant Principle [22], the DFE is globally asymptotically stable in Ω if $\mathcal{R}_0 \le 1$.

In contrast, if $\mathcal{R}_0 > 1$, then it follows from the continuity of the vector fields that $\frac{d\mathcal{L}_0}{dt} > 0$ in a neighborhood of the DFE in $\mathring{\Omega}$. Thus the DFE is unstable by the Lyapunov stability theory. The last part of the theorem can be proved by the persistent theory [23] which is similar to the proof of Theorem 2.5 in Gao and Ruan [24]. □

In addition, we have conducted an analysis on the global asymptotic stability of the endemic equilibrium [25, 26], and the details are presented in the following theorem. Essentially, these stability results establish $\mathcal{R}_0 = 1$ as a forward transcritical bifurcation point, or, a sharp threshold for disease dynamics, and indicate that reducing $\mathcal{R}_0$ to values at or below unity will be sufficient to eradicate the disease. In other words, our model (2.1) exhibits *regular* threshold dynamics. In order to simplify our notations, we will adopt the abbreviations

$$\beta_E = \beta_E(E), \ \beta_I = \beta_I(I), \ \beta_V = \beta_V(V), \ \beta_E^* = \beta_E(E_*), \ \beta_I^* = \beta_I(I_*), \ \beta_V^* = \beta_V(V_*).$$

Theorem 2.2. *Assume that $\beta_E(E)E$, $\beta_I(I)I$ and $\beta_V(V)V$ are non-decreasing functions of variables E, I and V, respectively. If $\mathcal{R}_0 > 1$, then the unique endemic equilibrium X_* of system (2.1) is globally asymptotically stable in $\mathring{\Omega}$.*

Proof. We let $L(y) = \int_{y_*}^y \frac{x-y_*}{x}dx$ for $y > 0$, where $y_* > 0$, and y can be replaced by S, E, I, or V. Clearly, $L(y) \ge 0$ with the equality holding if and only if $y = y_*$. Differentiating the four functions $L(S)$, $L(E)$, $L(I)$, $L(V)$ along the solution of system (2.1) and using the equilibrium equations yield

$$\frac{dL(S)}{dt} = \frac{S - S_*}{S}\frac{dS}{dt} \le \frac{S - S_*}{S}(\beta_E^* S_* E_* - \beta_E S E + \beta_I^* S_* I_* - \beta_I S I + \beta_V^* S_* V_* - \beta_V S V)$$

$$= \beta_E^* E_* S_*\left(1 - \frac{S_*}{S} - \frac{\beta_E E S}{\beta_E^* E_* S_*} + \frac{\beta_E E}{\beta_E^* E_*}\right) + \beta_I^* I_* S_*\left(1 - \frac{S_*}{S} - \frac{\beta_I I S}{\beta_I^* I_* S_*} + \frac{\beta_I I}{\beta_I^* I_*}\right)$$

$$+ \beta_V^* V_* S_*\left(1 - \frac{S_*}{S} - \frac{\beta_V V S}{\beta_V^* V_* S_*} + \frac{\beta_V V}{\beta_V^* V_*}\right),$$

$$\frac{dL(E)}{dt} = \frac{E - E_*}{E}\frac{dE}{dt} = \frac{E - E_*}{E}(\beta_E S E + \beta_I S I + \beta_V S V - (\alpha + \mu)E)$$

$$= \beta_E^* E_* S_* \left(\frac{\beta_E E S}{\beta_E^* E_* S_*} - \frac{E}{E_*} - \frac{\beta_E S}{\beta_E^* S_*} + 1\right) + \beta_I^* I_* S_* \left(\frac{\beta_I I S}{\beta_I^* I_* S_*} - \frac{E}{E_*} - \frac{\beta_I I S E_*}{\beta_I^* I_* S_* E} + 1\right)$$

$$+ \beta_V^* V_* S_* \left(\frac{\beta_V V S}{\beta_V^* V_* S_*} - \frac{E}{E_*} - \frac{\beta_V V S E_*}{\beta_V^* V_* S_* E} + 1\right).$$

Hence,

$$\frac{dL(S)}{dt} + \frac{dL(E)}{dt} \le \beta_E^* E_* S_* \left(2 - \frac{S_*}{S} - \frac{E}{E_*} + \frac{\beta_E E}{\beta_E^* E_*} - \frac{\beta_E S}{\beta_E^* S_*}\right)$$

$$+ \beta_I^* I_* S_* \left(2 - \frac{S_*}{S} - \frac{E}{E_*} + \frac{\beta_I I}{\beta_I^* I_*} - \frac{\beta_I I S E_*}{\beta_I^* I_* S_* E}\right)$$

$$+ \beta_V^* V_* S_* \left(2 - \frac{S_*}{S} - \frac{E}{E_*} + \frac{\beta_V V}{\beta_V^* V_*} - \frac{\beta_V V S E_*}{\beta_V^* V_* S_* E}\right)$$

$$\le \beta_E^* E_* S_* \left(\frac{\beta_E E}{\beta_E^* E_*} - 1\right)\left(1 - \frac{\beta_E^*}{\beta_E}\right)$$

$$+ \beta_I^* I_* S_* \left\{\left(\frac{\beta_I I}{\beta_I^* I_*} - 1\right)\left(1 - \frac{\beta_I^*}{\beta_I}\right) + \frac{I}{I_*} - \frac{E}{E_*} - \ln\frac{I}{I_*} + \ln\frac{E}{E_*}\right\}$$

$$+ \beta_V^* V_* S_* \left\{\left(\frac{\beta_V V}{\beta_V^* V_*} - 1\right)\left(1 - \frac{\beta_V^*}{\beta_V}\right) + \frac{V}{V_*} - \frac{E}{E_*} - \ln\frac{V}{V_*} + \ln\frac{E}{E_*}\right\}$$

$$\le \beta_I^* I_* S_* \left(\frac{I}{I_*} - \frac{E}{E_*} - \ln\frac{I}{I_*} + \ln\frac{E}{E_*}\right) + \beta_V^* V_* S_* \left(\frac{V}{V_*} - \frac{E}{E_*} - \ln\frac{V}{V_*} + \ln\frac{E}{E_*}\right).$$

The last inequality follows from the assumptions that $\beta_P(P)$ and $\beta_P(P)P$, where P can represent E, I, or V, are non-increasing and non-decreasing functions of P, respectively. This implies

$$1 - \frac{\beta_P^*}{\beta_P} \le 0 \iff P^* \le P \iff \frac{\beta_P P}{\beta_P^* P_*} - 1 \ge 0.$$

Similarly, one can verify that

$$\frac{dL(I)}{dt} = \alpha E_* \left(\frac{E}{E_*} - \frac{I}{I_*} - \frac{I_* E}{I E_*} + 1\right) \le \alpha E_* \left(\frac{E}{E_*} - \frac{I}{I_*} + \ln\frac{I}{I_*} - \ln\frac{E}{E_*}\right),$$

$$\frac{dL(V)}{dt} = \xi_1 E_* \left(\frac{E}{E_*} - \frac{V}{V_*} - \frac{V_* E}{V E_*} + 1\right) + \xi_2 I_* \left(\frac{I}{I_*} - \frac{V}{V_*} - \frac{V_* I}{V I_*} + 1\right)$$

$$\le \xi_1 E_* \left(\frac{E}{E_*} - \frac{V}{V_*} + \ln\frac{V}{V_*} - \ln\frac{E}{E_*}\right) + \xi_2 I_* \left(\frac{I}{I_*} - \frac{V}{V_*} + \ln\frac{V}{V_*} - \ln\frac{I}{I_*}\right).$$

Let $c_1 = \frac{\beta_I^* I_* S_*}{\alpha E_*} + \frac{\xi_2 \beta_V^* V_* S_*}{(w_1\xi_1 + \alpha\xi_2)E_*}$ and $c_2 = \frac{w_1 \beta_V^* V_* S_*}{(w_1\xi_1 + \alpha\xi_2)E_*}$. Then, we claim that

$$\mathcal{L}_1 = L(S) + L(E) + c_1 L(I) + c_2 L(V)$$

is a Lyapunov function for system (2.1). Clearly, $\mathcal{L}_1 \geq 0$ and

$$
\begin{aligned}
\frac{d\mathcal{L}}{dt} \leq &(\beta_I^* I_* S_* + \beta_V^* V_* S_* - c_1 \alpha E_* - c_2 \xi_1 E_*)\left(\ln \frac{E}{E_*} - \frac{E}{E_*}\right) \\
&+ (\beta_I^* I_* S_* - c_1 \alpha E_* + c_2 \xi_2 I_*)\left(\frac{I}{I_*} - \ln \frac{I}{I_*}\right) \\
&+ (\beta_V^* V_* S_* - c_2(\xi_1 E_* + \xi_2 I_*))\left(\frac{V}{V_*} - \ln \frac{V}{V_*}\right) \\
=&0
\end{aligned}
$$

with the equality holding if and only if $(S, E, I, V) = (S_*, E_*, I_*, V_*)$. Thus, one can easily see that the largest invariant set where $\frac{d\mathcal{L}}{dt} = 0$ is the singleton $\{X_* = (S_*, E_*, I_*, R_*, V_*)\}$. Therefore, X_* is globally asymptotically stable in $\mathring{\Omega}$.

$\square$

3. Numerical results

We now apply our model to study the COVID-19 epidemic in the city of Wuhan. We use the outbreak data published daily by WHO and other sources [7, 27–30]. These data sets contain the daily reported new cases, cumulative cases, and disease-caused deaths for the city of Wuhan, as well as each province in China and all other countries that have reported COVID-19 infection.

To conduct the numerical simulation, we consider the following functions for the three transmission rates in our model:

$$
\beta_E(E) = \frac{\beta_{E0}}{1 + cE}, \qquad \beta_I(I) = \frac{\beta_{I0}}{1 + cI}, \qquad \beta_V(V) = \frac{\beta_{V0}}{1 + cV}, \tag{3.1}
$$

where β_{E0}, β_{I0} and β_{V0} (all positive constants) denote the maximum values of these transmission rates, and c is a positive coefficient providing adjustment to the (otherwise constant) transmission rates. For example, $\beta_I(I)$ achieves its maximum when $I = 0$, with $\beta_I(0) = \beta_{I0}$, and $\beta_I(I)$ decreases when I is increasing. Note also that $\beta_E(E)E$, $\beta_I(I)I$ and $\beta_V(V)V$ are non-decreasing functions which satisfy the assumption of Theorem 2.2.

We implement our model and conduct numerical simulation for an epidemic period starting from January 23, 2020, when the city of Wuhan was placed in quarantine, to February 10, 2020. According to the estimate made by the Chinese government, about 9 million people remain in Wuhan after January 23 and they are not allowed to move out of the city. Meanwhile, only a relatively small number of people (mainly public health professionals) travel into the city since its lockdown. Thus, the influx rate Λ in our model is only based on newborns in Wuhan. The values of the transmission constants β_{E0} and β_{I0} can be found in a recent study [13]. The incubation period of the infection ranges between 2–14 days, with a mean of 5–7 days [31], and we take the value of 7 days in our model. The average recovery period is about 15 days [31], and so we set the disease recovery rate as $\gamma = 1/15$ per day. Members of the coronavirus family can survive in the environment from a few hours to several days [19], and we take the value of 1 day which results in a virus removal rate $\sigma = 1$ per day. Additionally, since the Chinese government has been implementing a very strict isolation policy and intense medical care for all the confirmed cases, represented by I in our model, the chance of those infected individuals

spreading the coronavirus to the environment connected with the general public is very low, and so we assume the virus shedding rate from the infected individuals is zero; i.e., $\xi_2 = 0$. Note that our results in Theorem 2.1 and Theorem 2.2 still hold in this case since the contribution of the coronavirus to the environmental reservoir remains a positive number $w_1\xi_1$. These and other parameters, their values and sources are provided in Table 1. There are three parameters, however, that remain to be determined: the environment-to-human transmission constant β_{V0}, the transmission adjustment coefficient c, and the virus shedding rate ξ_1 by the exposed individuals. The values of these parameters are not available in the literature because the models published thus far have not considered the environmental component for the COVID-19 infection, and they have generally applied constant transmission rates which remain fixed in time.

Table 1. Definitions and values of model parameters.

Parameter	Definition	Estimated mean value	Source
Λ	Influx rate	271.23 per day	[30]
β_{E0}	Transmission constant between S and E	3.11×10^{-8}/person/day	[13]
β_{I0}	Transmission constant between S and I	0.62×10^{-8}/person/day	[13]
β_{V0}	Transmission constant between S and V	fitting by data	-
c	Transmission adjustment coefficient	fitting by data	-
μ	Natural death rate	3.01×10^{-5} per day	[30]
$1/\alpha$	Incubation period	7 days	[31]
w	Disease-induced death rate	0.01 per day	[30]
γ	Recovery rate	1/15 per day	[31]
σ	Removal rate of virus	1 per day	[19]
ξ_1	Virus shedding rate by exposed people	fitting by data	-
ξ_2	Virus shedding rate by infected people	0 per person per day per ml	-

To estimate the values for these three parameters, similar to [32], we fit our model to the daily reported infection data for Wuhan from January 23 to February 10 by using the standard least squares method. Based on reported data, the initial condition is set as $(S(0), E(0), I(0), R(0), V(0)) = (89985051000, 475, 10, 10000)$ [33]. Figure 1 shows the numbers of cumulative confirmed cases in Wuhan during this period versus our fitting curve. The parameter values and their 95% confidence intervals are presented in Table 2. The normalized mean square error (NMSE) for the data fitting is found as 0.0058.

Table 2. Parameter estimates from data fitting.

Parameter	Fitting value	95% Confidence Interval
ξ_1	2.30	(0, 23.114)
β_{V0}	1.03×10^{-8}	(0, 5.074×10^{-8})
c	1.01×10^{-4}	(0.585×10^{-4}, 1.426×10^{-4})

Based on the parameter values from data fitting, we are able to evaluate the basic reproduction number $\mathcal{R}_0 = 4.25$. Specifically, we find that

$$\mathcal{R}_1 = 1.959, \qquad \mathcal{R}_2 = 0.789, \qquad \mathcal{R}_3 = 1.497,$$

which quantify the infection risk from each of the three transmission routes. Among these three components, the largest one ($\mathcal{R}_1$) comes from the exposed-to-susceptible transmission, since exposed individuals show no symptoms and can easily spread the infection to other people with close contact, often in an unconscious manner. Meanwhile, the smallest component ($\mathcal{R}_2$) comes from the infected-to-susceptible transmission, possibly due to the strict isolation policy on the symptomatic infected individuals. In addition, we observe that $\mathcal{R}_3 = 1.497$, showing a significant contribution from the environmental reservoir toward the overall infection risk.
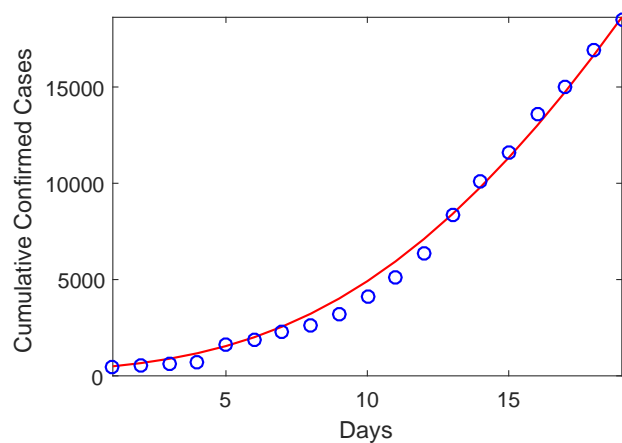

Figure 1. Cumulative confirmed cases for the city of Wuhan from January 23, 2020 to February 10, 2020. Circles (in blue) denote the reported cases and solid line (in red) denotes the simulation result. The basic reproduction number is $\mathcal{R}_0 = 4.25$ based on the parameters from Table 1 and the result of data fitting.
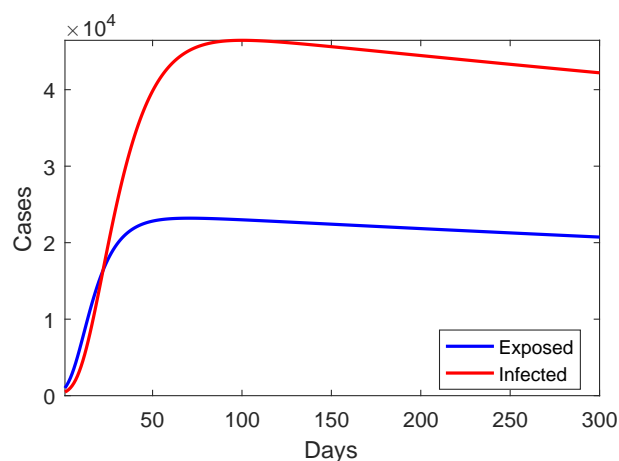

Figure 2. A simulation result for the outbreak size in Wuhan using the transmission rates formulated in Eq (3.1), the parameters from Table 1, and the result of data fitting.

Figure 2 displays a short-term prediction for I (the infected individuals) and E (the exposed individuals) in Wuhan using our model. It shows that the infection level, starting from January 23 (marked as day 0 in our simulation), would continue increasing for about 80 days, reach a peak value around 45,000 infections, and then gradually go down afterwards. Meanwhile, the long-term behavior of the epidemic would be determined by the property of the endemic equlibrium of the system, which is found as $X_* = (2583683, 1353, 2735, 6528015, 3111)$. A phase portrait of I vs. E is provided in Figure 3, where all the solution orbits converge to the endemic equlibrium, illustrating its global asymptotical stability that is stated in Theorem 2.2.
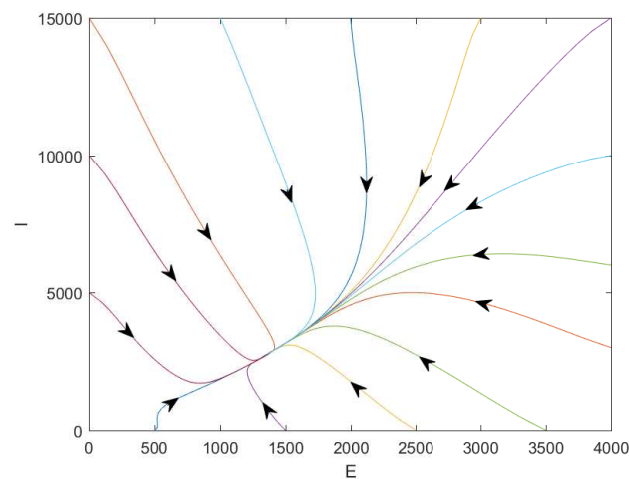

Figure 3. A typical phase portrait for I vs. E with $\mathcal{R}_0 = 4.25$. Every orbit starts with a different initial point and converges to the endemic equilibrium where $(E_*, I_*) = (1353, 2735)$.
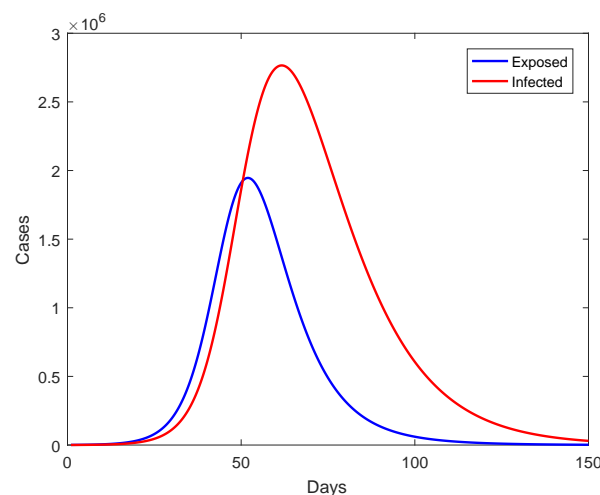

Figure 4. A simulation result for the outbreak size in Wuhan using the constant transmission rates given in Eq (3.2), the parameters from Table 1, and the result of data fitting.

In addition, we have performed a numerical test using simple, constant transmission rates in our model:

$$\beta_E(E) = \beta_{E0}, \qquad \beta_I(I) = \beta_{I0}, \qquad \beta_V(V) = \beta_{V0}, \qquad (3.2)$$

equivalent to setting $c = 0$ in Eq (3.1). This leaves two parameters, β_{V0} and ξ_1, to be estimated by data fitting. Using the same set of data, we find that $\xi_1 \approx 4.28$ with the 95% confidence interval $(0, 15.611)$, and $\beta_{V0} \approx 4.91 \times 10^{-10}$ with the 95% confidence interval $(4.218 \times 10^{-10}, 5.603 \times 10^{-10})$. The normalized mean square error (NMSE) for the data fitting is 0.0266, larger than that in the previous scenario, 0.0058. Meanwhile, Figure 4 shows a prediction of the Wuhan coronavirus outbreak size in this setting. Compared to Figure 2, We now clearly observe a significantly higher level of infection; particularly, the peak value appears at 2.8×10^6, which is extremely large and clearly unrealistic. The result demonstrates that using fixed transmission rates, which do not take into account the strong disease control measures currently on-going in Wuhan, may overestimate the epidemic severity and generate misguided information.

4. Discussion

We have proposed a mathematical model to investigate the on-going novel coronavirus epidemic in Wuhan, China. There are two unique features in our model: (1) the incorporation of an environmental reservoir into the disease transmission dynamics, and (2) the use of non-constant transmission rates which change with the epidemiological status and environmental conditions and which reflect the impact of the disease control measures implemented in Wuhan. We have conducted a detailed analysis of this model, and applied it to study the Wuhan epidemic using publicly reported data.

The basic reproduction number $\mathcal{R}_0$ of this model consists of three parts, representing the three different transmission routes; i.e., from the exposed individuals, the infected individuals, and the environmental reservoir, to the susceptible individuals. These three transmission modes collectively shape the overall disease risk of this epidemic, suggesting that intervention strategies should target all these three transmission routes. Our equilibrium analysis of this model shows that the disease dynamics exhibit a regular threshold at $\mathcal{R}_0 = 1$. We have established the global asymptotic stability of the disease-free equilibrium when $\mathcal{R}_0 < 1$, and the global asymptotic stability of the endemic equilibrium when $\mathcal{R}_0 > 1$.

Our numerical simulation results demonstrate the application of our model to the COVID-19 outbreak in Wuhan. Our model can fit the reported data well. Through data fitting, we obtain an estimate of basic reproduction number, $\mathcal{R}_0 = 4.25$. In particular, we find that the contribution of the environmental reservoir (measured by $\mathcal{R}_3$) is significant in shaping the overall disease risk. Our model predicts the appearance of an epidemic peak, after which the infection level would decrease and approach an endemic state in the long run. We also find that if we use constant transmission rates instead, the model would predict a much higher and unrealistic epidemic peak. This is caused by the fixed transmission rates that do not reflect the impact of on-going disease control measures. It is an indication that using epidemiologically and environmentally dependent transmission rates can potentially generate more practical simulation results.

At present, many aspects regarding the pathology, ecology and epidemiology of the novel coronavirus remain unknown, which adds challenges to the mathematical modeling. Particularly, in our current model, we have employed a bilinear incidence rate based on the law of mass action to represent the environment-to-human transmission route [34]. Practically, though, a saturation based incidence

rate might better characterize the environmental pathogen, and we hope to investigate it in our future modeling efforts. Meanwhile, the transmission rates β_E and β_V in our model depend on E and V, respectively, while in reality the exposed population E and the environmental pathogen concentration V may be unknown. To better quantify these transmission rates, we could instead assume that they are functions of I; i.e., $\beta_E(I)$ and $\beta_V(I)$, since the infected population I can be easily calculated from the reported data. Nevertheless, it is reasonable to assume that E and V are positively correlated to I, and so the qualitative properties of these transmission functions would remain the same under both formulations.

Given the current development of COVID-19, it is widely speculated that this disease would persist in the human world and become endemic. Our mathematical analysis and numerical simulation results support this speculation. The findings in this study imply that we should be prepared to fight the coronavirus infection for a much longer term than that of the current epidemic wave, in order to reduce the endemic burden and potentially eradicate the disease eventually. Among other intervention strategies, new vaccines for the novel coronavirus, which are currently in research and development, could play an important role in achieving that goal.

We emphasize that our data fitting is based on the reported confirmed cases in Wuhan from January 23 to February 10 in 2020. These confirmed cases were determined by the method of nucleic acid testing kits. On February 12, 2020, the national health commission in China started including cases confirmed by another method; i.e, clinical diagnosis, which refers to using CT imaging scans to diagnose patients. This change of criteria led to a surge of confirmed cases on February 12 (with an increase of about 14,000 new cases for Wuhan in a single day), and our current study does not take into account this factor. In this regard, our prediction of the epidemic duration and size should be interpreted as applicable only to the confirmed cases based on the previous, more strict, testing method. The issues regarding the accuracy, reliability and standard of reported data are complex and are beyond the scope of this work, which is more oriented on the mathematical modeling side. We plan to address the new development of the outbreak data in another piece of work in the near future. We also plan to expand our modeling efforts to the province and country levels beyond the epicenter, the city of Wuhan, and study the spread of the novel coronavirus in larger spatial scales.

Acknowledgments

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. *WHO statement regarding cluster of pneumonia cases in Wuhan, China.* 2020. Available from: `https://www.who.int/china/news/detail/09-01-2020-\` `who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china`.

2. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, et al., Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia, *N. Engl. J. Med.*, 2020.

3. P. Zhou, X. L. Yang, X. G. Wang, B. Hu, L. Zhang, W. Zhang, et al., Discovery of a novel coronavirus associated with the recent pneumonia outbreak in humans and its potential bat origin, *bioRxiv*, 2020.

4. L. E. Gralinski, V. D. Menachery, Return of the coronavirus: 2019-nCoV, *Viruses*, **12** (2020), 135.

5. T. Ellerin, The new coronavirus: What we do- and don't- know, *Harvard Health Blog*, January 25, 2020. Available from: `https://www.health.harvard.edu/blog/` `the-new-coronavirus-what-we-do-and-dont-know-2020012518747`.

6. *Centers for Disease Control and Prevention: 2019 novel coronavirus.* Available from: `https://www.cdc.gov/coronavirus/2019-ncov`.

7. *WHO novel coronavirus (2019-nCoV) situation reports.* Available from: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports`.

8. V. J. Munster, M. Koopmans, N. V. Doremalen, D. V. Riel, E. D. Wit, A novel coronavirus emerging in China – key questions for impact assessment, *N. Engl. J. Med.*, **382** (2020), 692–694.

9. *WHO statement regarding the outbreak of novel coronavirus (2019-nCoV)*, 2020. Available from: `https://www.who.int/news-room/detail/30-01-2020-statement-on-the-\` `second-meeting-of-the-international-health-regulations-(2005)-emergency-\` `committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov)`.

10. C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, et al., Transmission of 2019-nCoV Infection from an asymptomatic contact in Germany, *N. Engl. J. Med.*, 2020.

11. J.-T. Wu, K. Leung, G.M Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, *Lancet*, **395** (2020), 689–697.

12. J. M. Read, J. R. E. Bridgen, D. A. T. Cummings, A. Ho, C. P. Jewell, Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions, *medRxiv*, 2020.

13. B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, et al., Estimation of the transmission risk of 2019-nCoV and its implication for public health interventions, *J. Clin. Med.*, **9** (2020), 462.

14. N. Imai, A. Cori, I. Dorigatti, M. Baguelin, C. A. Donnelly, S. Riley, et al., Report 3: Transmissibility of 2019-nCoV, *Reference Source*, 2020. Available from: `https://www.imperial.ac.uk/` `mrc-global-infectious-disease-analysis/news--wuhan-coronavirus/`.

15. H. Zhu, Q. Gao, M. Li, C. Wang, Z. Feng, P. Wang, et al., Host and infectivity prediction of Wuhan 2019 novel coronavirus using deep learning algorithm, *bioRxiv*, 2020.

16. J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K. W. To, H. Chu, J. Yang, et al., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster, *Lancet*, **395** (2020), 514–523.

17. Z. J. Cheng, J. Shan, 2019 Novel coronavirus: where we are and what we know, *Infection*, (2020), 1–9.

18. A. Sahin, A. Erdogan, P. M. Agaoglu, Y. Dineri, A. Cakirci, M. Senel, et al., 2019 novel coronavirus (COVID-19) outbreak: A review of the current literature, *Eurasian J. Med. Oncol.*, **4** (2020), 1–7.

19. C. Geller, M. Varbanov, R. E. Duval, Human coronaviruses: Insights into environmental resistance and its influence on the development of new antiseptic strategies, *Viruses*, **4** (2012), 3044–3068.

20. G. Kampf, D. Todt, S. Pfaender, E. Steinmann, Persistence of coronaviruses on inanimate surfaces and its inactivation with biocidal agents, *J. Hosp. Infect.*, **104** (2020), 246–251.

21. P. V. D. Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48.

22. J. P. LaSalle, *The Stability of Dynamical Systems*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1976.

23. H. R. Thieme, Persistence under relaxed point-dissipativity (with application to an endemic model), *SIAM J. Math. Anal.*, **24** (1993), 407–435.

24. D. Gao, S. Ruan, An SIS patch model with variable transmission coefficients, *Math. Biosci.*, **232** (2011), 110–115.

25. M. Y. Li, J. S. Muldowney, P. V. D. Driessche, Global stability of SEIRS models in epidemiology, *Can. Appl. Math. Q.*, **7** (1999), 409–425.

26. Z. Shuai, P. V. D. Driessche, Global stability of infectious disease models using Lyapunov functions, *SIAM J. Appl. Math.*, **73** (2013), 1513–1532.

27. *Wikipedia on the timeline of the 2019–20 Wuhan coronavirus outbreak*. Available from `https://en.wikipedia.org/wiki/Timeline_of_the_2019\OT1\textendash20_Wuhan_coronavirus_outbreak`.

28. *National Health Commission of China daily reports on novel coronavirus (in Chinese)*. Available from: `http://www.nhc.gov.cn/yjb/pzhgli/new_list.shtml`.

29. *Sina news real-time reports on novel coronavirus (in Chinese)*. Available from: `https://news.sina.cn/zt_d/yiqing0121`.

30. *The government of Wuhan homepage*. Available from: `http://english.wh.gov.cn/`.

31. J. A. Spencer, D. P. Shutt, S. K. Moser, H. Clegg, H. J. Wearing, H. Mukundan, et al., Epidemiological parameter review and comparative dynamics of influenza, respiratory syncytial virus, rhinovirus, human coronvirus, and adenovirus, *m*edRxiv, 2020.

32. C. Yang, J. Wang, A cholera transmission model incorporating the impact of medical resources, *Math. Biosci. Eng.*, **16** (2019), 5226–5246.

33. *The Health Commission of Hubei Province*. Available from: `http://wjw.hubei.gov.cn/fbjd/dtyw/index_16.shtml`.

34. J. H. Tien, D. J. D. Earn, Multiple transmission pathways and disease dynamics in a waterborne pathogen model, *Bull. Math. Biol.*, **72** (2010), 1506–1533.