Analysis of cholera epidemics with bacterial growth and spatial movement

Xueying Wang\textsuperscript{a*} and Jin Wang\textsuperscript{b}

\textsuperscript{a}Department of Mathematics, Washington State University, Pullman, WA 99164, USA; \textsuperscript{b}Department of Mathematics, University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA

(Received 26 February 2014; accepted 3 October 2014)

In this work, we propose novel epidemic models (named, susceptible–infected–recovered–susceptible–bacteria) for cholera dynamics by incorporating a general formulation of bacteria growth and spatial variation. In the first part, a generalized ordinary differential equation (ODE) model is presented and it is found that bacterial growth contributes to the increase in the basic reproduction number, $R_0$. With the derived basic reproduction number, we analyse the local and global dynamics of the model. Particularly, we give a rigorous proof on the endemic global stability by employing the geometric approach. In the second part, we extend the ODE model to a partial differential equation (PDE) model with the inclusion of diffusion to capture the movement of human hosts and bacteria in a heterogeneous environment. The disease threshold of this PDE model is studied again by using the basic reproduction number. The results on the threshold dynamics of the ODE and PDE models are compared, and verified through numerical simulation. Additionally, our analysis shows that incorporating diffusive spatial spread does not produce a Turing instability when $R_0$ associated with the ODE model is less than the unity.

Keywords: cholera model; basic reproduction number; global asymptotic stability; geometric approach; disease threshold dynamics

1. Introduction

Recent years witnessed an increasing number of cholera outbreaks worldwide [41], including one of the largest cholera epidemics in modern history that took place in Haiti during 2010–2012 with more than 530,000 reported cases and over 7000 deaths, and the worst African cholera outbreak in the past 20 years that devastated Zimbabwe during 2008–2009 with nearly 100,000 reported cases and more than 4000 deaths. These outbreaks, with their increased frequency and severity, indicate that our current knowledge in cholera dynamics and public health guidelines to control the disease is not adequate.

Cholera is an acute intestinal infectious disease caused by the bacterium \textit{Vibrio cholerae}. It can spread rapidly and lead to death within days if left untreated. The complexity of cholera dynamics stems from the fact that both direct (i.e. human-to-human) and indirect (i.e. environment-to-human) routes are involved in the disease transmission [22]. Thus, the dynamics of cholera involve multiple interactions among the human host, the pathogen, and the environment [26].

\*Corresponding author. Email: xueying@math.wsu.edu
Author Email: jin-wang02@utc.edu

© 2014 The Author(s). Published by Taylor & Francis.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.
In an effort to gain deeper understanding of cholera dynamics, many mathematical models have recently been proposed and analysed. Codeço published a model [7] that added the pathogen concentration in the water supply into a regular susceptible–infected–recovered (SIR) epidemiological model. Hartley et al. [13] included a hyperinfectious state of the pathogen, representing the ‘explosive’ infectivity of freshly shed vibrios [1]. Tien and Earn [35] and Mukandavire et al. [22] explicitly considered both human-to-human and environment-to-human transmission pathways. Wang and Liao [38] proposed a cholera model that incorporated general incidence and pathogen functions. Shuai et al. [32] investigated cholera dynamics with both hyperinfectivity and temporary immunity. Other cholera models include, but not limited to, [2–4,6,18,21,28,31,34,36].

One limitation in most of these mathematical cholera studies is that the dynamics of the pathogen (i.e. the vibrios) are poorly addressed. A standard assumption in almost all the cholera models is that the vibrios cannot sustain themselves in the absence of human contribution (e.g. shedding from infected individuals and inflow from contaminated sewage). This assumption allows a simple, often linear, representation of the rate of change for the bacterial density: a positive contribution from the infected human population, and a negative contribution due to natural death of the vibrios. On the other hand, there have been strong evidences [8] that the vibrios can independently persist in the environment and, consequently, their intrinsic growth and decay may play an essential role in shaping cholera epidemics. Only very few studies [3,18] so far have taken into account the intrinsic bacterial dynamics in cholera modelling.

Another challenge in current mathematical cholera studies is that spatial heterogeneity is rarely considered, resulting in insufficient understanding of the spatial spread of cholera infection. In the work of [22], basic reproduction numbers were estimated for all the 10 provinces in Zimbabwe and the results were highly heterogeneous, implying that the underlying transmission pattern varied widely throughout the country. Similarly, Tuite et al. [36] obtained very different reproduction numbers for the 10 administrative departments in Haiti during the recent cholera outbreak. The findings in these studies underscore the importance of spatial heterogeneity in cholera transmission and the design of control strategies, and suggest that more work is demanded towards better understanding of this issue. In addition, Bertuzzo et al. [4] developed a simple partial differential equation (PDE) model, building on the framework of Codeço’s model [7], to investigate the spatial movement of the pathogen in cholera epidemic setting.

In the present paper, we aim to address the aforementioned challenges in cholera modelling by proposing new models that incorporate both intrinsic bacterial dynamics and spatial variation. We will start with an ordinary differential equation (ODE) model that allows a general representation of the direct and indirect transmission pathways and a general description of the bacterial growth. We then derive the basic reproduction number of this model and analyse the local and global stabilities for the disease-free equilibrium (DFE) and endemic equilibrium (EE). Particularly, we will apply the geometric approach originally introduced by Li and Muldowney [20] to the endemic global stability analysis. Next, we extend this model to a PDE system by adding diffusion terms for both the human population and the pathogen. We investigate the threshold dynamics of this PDE model by analysing and estimating its basic reproduction number. To our knowledge, no prior work has been published on the threshold dynamics of PDE cholera models. In addition, we verify our analysis using numerical simulation results.

The remainder of this paper is organized as follows. In Section 2, the ODE model is presented and necessary assumptions are stated, followed by a careful analysis of the threshold dynamics based on the derived basic reproduction number. In Section 3, the PDE model is introduced, the possibility of Turing instability is investigated, and the basic reproduction number is analysed and compared with that of the ODE model. In Section 4, numerical results are presented to validate the analytical predictions. Finally, conclusions are drawn and some discussion is presented in Section 5.
2. A generalized SIRS-B cholera epidemic model

2.1. Model description

Cholera infection consists of two populations: human hosts and bacteria. Since cholera infection does not invoke a long-lasting immune response [19,24,43], we apply the standard susceptible–infected–recovered–susceptible (SIRS) epidemic framework for the infection with the host population. Moreover, we assume that (a) a susceptible host becomes infected either by direct contact with infectious hosts or via indirect contact with bacteria in contaminated water; (b) infectious hosts contaminate the environment through shedding of the bacteria. A compartmental diagram of a generalized SIRS-B epidemic model (where B stands for the bacterial concentration in the environment) describing the dynamics of cholera infection is displayed in Figure 1. This generalized model can be written as

\[
\begin{align*}
\frac{dS}{dt} &= b - S f_1(I) - S f_2(B) - dS + \sigma R, \\
\frac{dI}{dt} &= S f_1(I) + S f_2(I) - (d + \gamma) I, \\
\frac{dR}{dt} &= \gamma I - (d + \sigma) R, \\
\frac{dB}{dt} &= \xi I + h(B) - \delta B.
\end{align*}
\]  

(1)

Here \(S, I\) and \(R\) represent the number of SIR hosts, respectively, and \(B\) is the concentration of the bacteria (vibrios) in the contaminated water. The parameter \(b\) describes the influx (or, recruitment) of susceptible hosts. The functions \(f_1(I)\) and \(f_2(B)\) depict the direct and indirect transmission rates, respectively. For example, \(f_1(I) = 0\) and \(f_2(B) = aB/(B + \kappa)\) in the model of Codeco [7] (where \(a\) is the contact rate with contaminated water, and \(\kappa\) is the half saturation rate that describes the infectious dose in water sufficient to produce disease in 50% of those exposed), \(f_1(I) = \beta I\) and \(f_2(B) = \lambda B\) in the model proposed by Gosh et al. [11] (where \(\beta\) and \(\lambda\) represent the direct and indirect transmission parameters due to the human-to-human and the environment-to-human interactions, respectively), and \(f_1(I) = \beta_h I\) and \(f_2(B) = \beta_e B/(B + \kappa)\) in the model of Mukandavire et al. [22] (where \(\beta_h\) and \(\beta_e\) represent the direct and indirect transmission parameters). In addition, \(d\) is the natural death rate of each host class, \(\gamma\) is the recovery
rate of infectious individuals, \( \sigma \) denotes the rate at which recovered individuals lose immunity, \( \xi \) is the shedding rate of bacteria by infectious hosts, and \( \delta \) is the natural death rate of the bacteria. All these aforementioned model parameters are assumed to be positive.

Furthermore, as most existing cholera models (e.g. \([7,13,17,18,22]\)), disease-induced mortality is assumed to be negligible. A reason for making this assumption is that, with a few exceptions, the death rate for cholera is generally quite low (about 1%). For example, WHO states that ‘In 2012, the overall case fatality rate for cholera was 1.2%’ \([40]\).

Finally, the function \( h(B) \) is introduced to describe the intrinsic growth of the bacteria. The functions \( f_1, f_2 \) and \( h \) are assumed to be all differentiable in this study.

We assume that

(H1) \( (1) f_1(I) \geq 0; (2) f_1(0) = 0; (3) f_1'(I) > 0; (4) f_1''(I) \leq 0. \)

(H2) \( (1) f_2(B) \geq 0; (2) f_2(0) = 0; (3) f_2'(B) > 0; (4) f_2''(B) \leq 0. \)

(H3) \( (1) h(0) = 0; (2) h'(B) \leq 0. \)

Biologically, the assumptions (H1) and (H2) state that the disease transmission rates are monotonically increasing, but also subject to saturation effects. The incidence functions employed in most of the existing cholera models (e.g. \([7,11,22,35]\)) satisfy these conditions. Meanwhile, the assumption (H3) states that the bacterial growth rate is also subject to saturation effects. We mention, however, that (H3) excludes some more complicated growth models such as the logistic growth with a threshold \( T_B \) (i.e. \( h(B) = rB(1 - B/K)(B/T_B - 1) \)) and extended logistic growth models \([25]\).

2.2. Basic reproduction number

In epidemic models, one of the main concerns is to quantify the infection risk so as to effectively control the disease. An important disease control threshold is the basic reproduction number, \( R_0 \), which measures the expected number of secondary infections caused by one infectious individual during its infectious period in an otherwise susceptible population \([9,10]\). It can be computed using the next-generation matrix theory \([10]\). We will focus our attention on the basic reproduction number in this paper, but see Appendix 1 for two other types of threshold quantities which are closely related to \( R_0 \).

Let \( N^* = b/d \). One can easily see from direct calculation that the DFE is given by \((N^*, 0, 0, 0)\). The Jacobian matrix \( J \) at the DFE is decomposed as follows: \( J = F - V \), where \( F \) is the matrix characterizing the generation of secondary infectious cases/agents, and \( V \) is the matrix of transition rates between compartments. We assume that bacteria population in the environment is regarded as a reservoir of the infection. The free-living bacteria adapt to the environment through physiological and genetic changes that can promote their survival and growth \([8]\). Thus, we assume that the bacteria can possibly be self-maintained through growth in the environment, and secondary bacteria can be introduced through pathogen shedding by infectious humans. For simplicity of notations, we write

\[
f(I,B) = f_1(I) + f_2(B), \quad f_I = f_1'(0), \quad f_B = f_2'(0) \quad \text{and} \quad g = h'(0).
\]

Based on the standard next-generation matrix technique \([10]\) and our assumptions, matrices \( F \) and \( V \) have the following forms:

\[
F = \begin{bmatrix} N^* f_I & N^* f_B \\ \xi & g \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} (d + \gamma) & 0 \\ 0 & \delta \end{bmatrix}.
\]
Hence, the next-generation matrix is

\[
K = FV^{-1} = \begin{bmatrix}
\frac{N^*f_I}{d + \gamma} & \frac{N^*f_B}{\delta} \\
\frac{\xi}{d + \gamma} & \frac{g}{\delta}
\end{bmatrix}.
\] (3)

The basic reproduction number \( R_0 \) is given by

\[
R_0 = \rho(K) = \frac{1}{2} \left[ \frac{N^*f_I}{d + \gamma} + \frac{g}{\delta} + \sqrt{\left(\frac{N^*f_I}{d + \gamma} - \frac{g}{\delta}\right)^2 + 4\frac{\xi N^*f_B}{\delta(d + \gamma)}} \right].
\] (4)

Here \( \rho(M) \) denotes the spectral radius for any \( n \times n \) matrix \( M \).

In particular, the quantity \( R_0^h := \frac{N^*f_I}{(d + \gamma)} \) (resp. \( R_0^e := \frac{g}{\delta} \)) represents the average number of secondary infections through human-to-human (resp. the environment-to-human) transmission produced by one infectious host during its infectious period. \( g/\delta \) depicts the average number of secondary free-living bacteria caused by one bacterium during its lifetime in the environment.

Recall that \( \xi > 0 \) by the model assumption. It is straightforward to see from assumptions (H1)–(H3) that

\[ R_0 > \max\{R_0^h, R_0^e\}. \]

This result shows that (a) the cholera infection with bacterial presence is more severe than that in the absence of bacteria, which indicates that if an outbreak occurs in human population without bacteria \( (R_0^h > 1) \), then an outbreak will certainly take place in the presence of bacteria \( (R_0 > 1) \); (b) if the free-living bacteria can sustain themselves in the environment without hosts \( (R_0^e > 1) \), then disease will persist in the presence of hosts \( (R_0 > 1) \).

Suppose that all the parameters except \( g \) are constants. Consider \( R_0 \) as a function of \( g \), that is, \( R_0 = R_0(g) \). Based on our model assumptions, one can verify by direct calculation that

\[ R_0(g) > R_0(0) \quad \text{if} \quad g > 0. \]

This indicates that the basic reproduction number of Equation (1) in the presence of bacterial growth is higher than that without inclusion of bacterial growth. In other words, the persistence and intrinsic growth of vibrios would increase the risk of cholera epidemics. Moreover, it is easy to see that \( R_0 > 1 \) if \( g \geq \delta \). This implies that with strong growth and prevalence of vibrios, utilizing only vaccination and/or antibiotics would not be able to reduce \( R_0 \) below 1. Instead, improving infrastructure and providing clear water would be of fundamental importance in controlling cholera in the long run.

2.3. Equilibrium analysis

It follows by direct calculation that an equilibrium for the system (1) is a solution of the following equations:

\[
S + I + R = N^*,
\] (5)

\[
R = \frac{\gamma}{d + \sigma} I,
\] (6)

\[
I = \phi(B) := \frac{-h(B) + \delta B}{\xi},
\] (7)

\[
Sf_1(I) + Sf_2(B) = (d + \gamma)I.
\] (8)
By Equations (5) and (6), \( S := S(I) = N^* - (1 + \gamma/(d + \sigma))I \). By assumption (H2), \( f'_2(B) > 0 \), and hence there exists a unique function \( f^{-1}_2 \), such that \( f^{-1}_2(f_2(B)) = B \). Let \( \alpha = 1 + \gamma/(d + \sigma) \). In view of Equation (8),

\[
B = \psi(I) := f^{-1}_2 \left( \frac{(d + \gamma)I}{N^* - \alpha I} - f_1(I) \right).
\]

The intersection points of the curves \( I = \phi(B) \) and \( B = \psi(I) \) in \( \mathbb{R}^2_+ \) determine the equilibria. First, let us consider the concavity of the curves (7) and (9) on the \( B - I \) plane. By (H3),

\[
\phi''(B) = -\frac{h''(B)}{\xi} \geq 0
\]

and \( \phi(B) \) is concave up. On the other hand, by a direct calculation, we have

\[
\psi''(I) = \left[ \frac{2\alpha(d + \gamma)N^*}{S^3} - f''_1(I) - f''_2(B)(\psi'(I))^2 \right] / f'_2(B).
\]

By assumptions (H1) and (H2), Equation (10) yields \( \psi''(I) > 0 \), and hence \( \psi(I) \) is concave up. So, there exist at most two points of intersection in \( \mathbb{R}^2_+ \). Notice that \( \phi(0) = \psi(0) = 0 \). There always exists a unique DFE, \( (I, B) = (0, 0) \), where \( S = N^* \) and \( R = 0 \). There may be a second, EE, where the two curves intersect in the interior of \( \mathbb{R}^2_+ \). This depends on the slopes of the curves at zero: \( \psi'(0) = [(d + \gamma)/N^* - f'_1(0)]/f'_2(0) \) and \( \phi'(0) = (-g + \delta)/\xi \). There are two cases where the two curves intersect in the interior of \( \mathbb{R}^2_+ \), leading to an EE: (a) \( g \geq \delta \) or (b) \( g < \delta \) and \( \psi'(0) < [\phi'(0)]^{-1} \). The stability of the equilibrium will be established in Theorem 2.1.

First, let us consider \( \rho(M) \) for a square matrix \( M \) of size 2. Denote ‘if and only if’ as \( \iff \).

**Lemma 2.1** Let \( M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \) with \( a + d \geq 0 \) and \( bc \geq 0 \). Then

(a) \( \rho(M) < 1 \iff a < 1, d < 1 \) and \( \text{trace}(M) < \det(M) + 1 \).

(b) \( \rho(M) = 1 \iff a \leq 1, d \leq 1 \) and \( \text{trace}(M) = \det(M) + 1 \).

(c) \( \rho(M) > 1 \iff \text{trace}(M) \geq 2, \) or \( \text{trace}(M) < 2 \) and \( \text{trace}(M) > \det(M) + 1 \).

**Proof** (a) By the assumption on \( M \), it is easy to verify that

\[
\rho(M) = \frac{1}{2}[a + d + \sqrt{(a - d)^2 + 4bc}].
\]

Thus,

\[
\rho(M) < 1 \iff \sqrt{(a - d)^2 + 4bc} < 2 - (a + d),
\]

\[
\iff (a - d)^2 + 4bc < (2 - (a + d))^2 \quad \text{and} \quad 2 - (a + d) > 0,
\]

\[
\iff a + d < ad - bc + 1 \quad \text{and} \quad (a + d) < 2,
\]

\[
\iff \text{trace}(M) < \det(M) + 1 \quad \text{and} \quad \text{trace}(M) < 2. \quad (12)
\]

By Equation (11), one can directly verify that \( \rho(M) \geq \max\{a, d\} \geq 1 \) if \( d \geq 1 \) or \( a \geq 1 \). Thus, \( \rho(M) < 1 \) implies \( a < 1 \) and \( d < 1 \). Together with Equation (12), we show that

\[
\rho(M) < 1 \iff a < 1, d < 1 \quad \text{and} \quad \text{trace}(M) < \det(M) + 1. \quad (13)
\]

The other direction of Equation (13) is trivial by virtue of Equation (12). Likewise, one can prove (b) and (c).
Remark of Lemma 2.1 (c): If \( bc > 0 \), one can verify that

\[
\rho(M) = 1 \iff a < 1, d < 1 \quad \text{and} \quad \text{trace}(M) = \det(M) + 1.
\]

\[
\rho(M) > 1 \iff a \geq 1, \text{or} \quad d \geq 1, \text{or} \quad a < 1, d < 1 \quad \text{and} \quad \text{trace}(M) > \det(M) + 1.
\]

Based on the expression of \( R_0 \) in Equation (4), we can easily obtain that when \( g \geq \delta, R_0 > 1 \).

Meanwhile, note that

\[
\phi'(B) = \frac{\delta - h'(B)}{\xi} \quad \text{and} \quad \psi'(I) = \frac{N^*(d + \gamma)/(N^* - \alpha I)^2 - f_I'(I)}{f_2'(B)},
\]

which yields

\[
[\phi'(0)]^{-1} = \frac{\xi}{\delta - g} \quad \text{and} \quad \psi'(0) = \frac{d + \gamma - N^* f_I}{N^* f_B}.
\]

**Lemma 2.2** The following statements hold:

1. \( R_0 \leq 1 \iff g < \delta \) and \( \psi'(0) \geq [\phi'(0)]^{-1} \).
2. \( R_0 > 1 \iff g \geq \delta, \) or \( g < \delta, \) and \( \psi'(0) < [\phi'(0)]^{-1} \).

**Proof** (1) Applying Lemma 2.1(a), and Equation (14) to \( K \), we obtain

\[
R_0 \leq 1 \iff N^* f_I < d + \gamma, \ g < \delta, \quad \text{and} \quad N^* f_B \xi \leq (d + \gamma - N^* f_I)(\delta - g),
\]

\[
\iff g < \delta \quad \text{and} \quad \frac{d + \gamma - N^* f_I}{N^* f_B} \geq \frac{\xi}{\delta - g},
\]

\[
\iff g < \delta \quad \text{and} \quad \psi'(0) \geq [\phi'(0)]^{-1}.
\]

This proves (1). Similarly, we can show (2).

We now prove the following theorem regarding the local stabilities of the DFE and EE of the system (1). The local stability of the DFE; that is, the DFE is locally asymptotically stable for \( R_0 < 1 \) and unstable for \( R_0 > 1 \), can be obtained directly from Theorem 2 of [10]. To demonstrate the local stability of the DFE when \( R_0 = 1 \), however, we provide a proof below which automatically integrates the cases for \( R_0 < 1 \) and \( R_0 = 1 \).

**Theorem 2.1** (1) If \( R_0 \leq 1 \), then Equation (1) has only the DFE and it is locally asymptotically stable.

(2) If \( R_0 > 1 \), then Equation (1) has two equilibria: the DFE and the EE. Furthermore, the DFE is unstable and the EE is locally asymptotically stable.

**Proof** Linearizing (1) at the equilibrium \((S, I, R, B)\), we obtain the Jacobian matrix

\[
J = \begin{bmatrix}
-(f_I(I) + f_2(B)) - d & \frac{-S^* f_I(I)}{f_2'(B)} & \sigma & \frac{-S^* f_2'(B)}{f_2'(B)} \\
\frac{f_I(I) + f_2(B)}{f_2'(B)} & \frac{S^* f_I(I) - (d + \gamma)}{f_2'(B)} & 0 & \frac{S^* f_2'(B)}{f_2'(B)} \\
0 & \gamma & -(d + \sigma) & 0 \\
0 & \xi & 0 & h'(B) - \delta
\end{bmatrix}.
\]

**Case 1**: \( R_0 \leq 1 \). By Lemma 2.2, \( R_0 \leq 1 \) implies \( g < \delta \) and \( \psi'(0) \geq [\phi'(0)]^{-1} \). Hence we see that Equation (1) has only one equilibrium and it is the DFE. Evaluating \( J \) at the DFE, we find that \( \lambda_1 = -d, \lambda_2 = -(d + \sigma), \lambda_3 + \lambda_4 = (N^* f_I - (d + \gamma)) + (g - \delta) \) and \( \lambda_3 \lambda_4 = (N^* f_I - (d + \gamma))(g - \delta) - N^* \xi f_B \). Clearly, \( \lambda_1 \) and \( \lambda_2 \) are both negative. Hence the local stability of the
DFE relies on the sign of $\lambda_3$ and $\lambda_4$. In view of $\psi'(0) \geq [\phi'(0)]^{-1} = \xi/(\delta - g) > 0$, we have $\psi'(0) = [(d + \gamma)/N^* - f_I]/f_B > 0$, $f_B > 0$ implies $N^* f_I - (d + \gamma) < 0$. Together with $g < \delta$, it gives $\lambda_3 + \lambda_4 < 0$. By $[\phi'(0)]^{-1} < \psi'(0)$ and $g < \delta$, one can directly verify that $\lambda_3 \lambda_4 > 0$. This implies $\lambda_3 < 0$ and $\lambda_4 < 0$. Therefore, the DFE is locally asymptotically stable.

Case 2: $R_0 > 1$. In view of Lemma 2.2, $R_0 > 1$ if and only if (a) $g \geq \delta$; (b) $g < \delta$ and $\psi'(0) < [\phi'(0)]^{-1}$. In either case, Equation (1) has the DFE and EE. By a similar argument as that of Case 1, one can verify that the DFE is unstable. It remains to show that the EE is locally asymptotically stable.

Let $N(t) = (S + I + R)(t)$. We consider an equivalent system of Equation (1)

$$
\frac{dN}{dt} = b - dN,
$$

$$
\frac{dI}{dt} = (N - I - R)(f_I(I) + f_2(B)) - (d + \gamma)I,
$$

$$
\frac{dR}{dt} = \gamma I - (d + \sigma)R,
$$

$$
\frac{dB}{dt} = \xi I + h(B) - \delta B.
$$

Then the Jacobian matrix of Equation (17) evaluated at the EE $(N^*, I, R, B)$ is given by

$$
\hat{J} = \begin{bmatrix}
-d & 0 & 0 & 0 \\
-f(I, B) & -f(I, B) + Sf'_I(I) - (d + \gamma) & -f(I, B) + Sf'_2(B) & 0 \\
0 & \gamma & -(d + \sigma) & 0 \\
0 & \xi & 0 & h'(B) - \delta
\end{bmatrix},
$$

where $f(I, B) = f_I(I) + f_2(B)$, and $S = N^* - I - R$. It is easy to verify that the characteristic polynomial of $\hat{J}$ is

$$
\det(\lambda I - \hat{J}) = (\lambda + d)(\lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3),
$$

where

$$
a_1 = -(c_1 + c_3 + c_4),
$$

$$
a_2 = c_1 c_4 + c_3(c_1 + c_4) + c_0 + \xi c_2,
$$

$$
a_3 = -c_3(c_1 c_4 + c_0) - \xi c_2 c_4,
$$

with

$$
c_0 = f(I, B) \gamma,
$$

$$
c_1 = -f(I, B) + Sf'_I(I) - (d + \gamma),
$$

$$
c_2 = -Sf'_2(B),
$$

$$
c_3 = h'(B) - \delta,
$$

$$
c_4 = -(d + \sigma).
$$

Clearly, $\lambda_1 = -d < 0$. Thus, the stability of the EE is determined by the zeros of

$$
\lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3 = 0.
$$

It follows from the Routh–Hurwitz criterion that to verify that the EE is stable, it suffices to show that

$$
a_1 > 0, a_2 > 0, a_3 > 0, a_1 a_2 > a_3,
$$

(21)
Thus, by virtue of

\[ a_2 = c_1c_4 + c_0 + c_3c_4 + (c_1c_3 + \xi c_2), \]
\[ a_3 = (-c_4)(c_1c_3 + \xi c_2) + (-c_3)c_0, \]
\[ a_1a_2 - a_3 = -(c_1 + c_3)((c_1c_3 + \xi c_2) + c_4(c_3 + c_4) + c_1c_4) - (c_1 + c_4)c_0. \]

We now claim that

\begin{align*}
(1) & \quad c_0 \geq 0, c_i < 0 \ (i = 2, 3, 4), \\
(2) & \quad c_1c_3 + \xi c_2 > 0.
\end{align*}

Clearly, Equation (21) is satisfied if Equation (23) holds. It remains to verify Equation (23). In the following statement, \((S, I, R, B) \in \mathbb{R}_+^5\) is assumed to the EE. It is clear that \(c_0 \geq 0, c_2 < 0\) and \(c_4 < 0\) hold, since \(f(I, B) \geq 0, \gamma \geq 0, \sigma \geq 0, S > 0, f'_2(B) > 0\) and \(d > 0\). First, we want to show that \(Sf'_1(I) - (d + \gamma) \leq 0\). In view of \(I = 0, S(f_1(I) + f_2(B)) = (d + \gamma)I\). It follows from \(f_1(I) \geq f'_1(I)I\) that

\[ Sf'_1(I) - (d + \gamma) = \frac{S}{I}(f'_1(I)I - f_1(I) - f_2(B)) \leq 0. \]

Together with \(f(I, B) > 0\) implies \(c_1 < 0\). Second, we will show \(c_3 < 0\). Notice \(\delta = (\xi I + h(B))/B\). We find that

\[ h'(B) - \delta = h'(B) - \frac{(\xi I + h(B))}{B} = \frac{(h'(B)B - h(B) - \xi I)}{B} < 0, \]

by virtue of \(h(B) \geq h'(B)B\) and \(\xi > 0\). This gives \(c_3 < 0\). Last, we will verify the second inequality of Equation (23) holds. Notice that

\[ c_1c_3 + \xi c_2 = -f(I, B)(h'(B) - \delta) + ((h'(B) - \delta)(Sf'_1(I) - (d + \gamma)) - \xi Sf'_2(B)). \]

To show \(c_1c_3 + \xi c_2 > 0\), it suffices to show \((h'(B) - \delta)(Sf'_1(I) - (d + \gamma) - \xi Sf'_2(B)) \geq 0\). By \(h(B) \geq h'(B)B, \xi I + h(B) - \delta B = 0\) implies \(\xi I + (h'(B) - \delta)B \leq 0\). Thus,

\[ B_0 := B - \frac{\xi I}{(\delta - h'(B))} \geq 0. \]

On the other hand, a Taylor expansion of \(f_2\) at \(B\) yields \(0 \leq f_2(B_0) = f_2(B) + f'_2(B)(B_0 - B) + \frac{f''_2(\theta_B)(B_0 - B)^2}{2}, \) for some \(\theta_B\) depending on \(B_0\) and \(B\). By assumption (H2), we see that

\[ f_2(B) + f'_2(B)(B_0 - B) \geq 0 \]

and hence

\[ f_2(B) \geq f'_2(B)\frac{\xi I}{\delta - h'(B)}. \]

Together with \(f_1(I) \geq f'_1(I)I\) gives

\[ 0 = Sf_1(I) + Sf_2(B) - (d + \gamma)I \geq Sf'_1(I)I + Sf'_2(B)\frac{\xi I}{h'(B)} - (d + \gamma)I. \]

Thus,

\[ (h'(B) - \delta)(Sf'_1(I) - (d + \gamma)) - \xi Sf'_2(B) \geq 0. \]

The proof is complete. □

Building on the result in Theorem 2.1, we now proceed to analyse the global stabilities of the equilibria for our cholera model (1).
2.4. Global dynamics

Adding the first three equations of Equation (1), we get \( \frac{dN}{dt} = b - dN \). Hence \( 0 \leq N(t) \leq N^* \) if \( 0 \leq N(0) \leq N^* \). Now consider the growth of bacteria, which is described by \( h(B) \). We introduce another regulation on the intrinsic bacterial growth:

\[(H4) \limsup_{B \to \infty} h(B) \leq (-r + \delta)B \text{ for some positive constant } r.\]

Note that the rate of change of the bacterial concentration due to the intrinsic growth and death is \( h(B) - \delta B \). Thus, biologically, the assumption \( (H4) \) implies that in the absence of external contribution (e.g., shedding from infected people), the bacteria would eventually decay (at least in a linear manner) when the bacterial population size is large. Based on this assumption, there exists \( M > 0 \) such that \( 0 \leq B(t) \leq M \) as long as \( 0 \leq B(0) \leq M \).

Define the domain

\[ \Delta = \{(S, I, R, B) : S, I, R, B \geq 0, S + I + R \leq N^*, B \leq M\}. \]

Clearly, if the solution of system (1) is initially in \( \Delta \), it will remain in \( \Delta \). Hence, \( \Delta \) is positively invariant for model (1). The following theorem summarizes the global dynamics of model (1).

**Theorem 2.2** Suppose that assumptions \((H1)-(H4)\) hold.

1. If \( R_0 < 1 \), then the deterministic model (1) has a unique DFE that is globally asymptotically stable in \( \Delta \).
2. If \( R_0 > 1 \), then the EE of Equation (1) is globally asymptotically stable in the interior of \( \Delta \), provided \( b = dN \) and \( \sup_{(S,I,B)\in\Delta} \{2(Sf_1'(I)) + \sigma\} \leq \gamma. \)

**Proof** Case (1): \( R_0 < 1 \).

Then Theorem 2.1 implies that the system (1) has a unique DFE, and it is locally stable. By assumptions \((H1)-(H3)\), \( f_1(I) \leq f_1I, f_2(B) \leq f_2B \) and \( h(B) \leq gB \) for all \( I \) and \( B \). Hence, Equation (1) implies

\[
\begin{align*}
\frac{dI}{dt} &\leq N^* (f_1I + f_2B) - (d + \gamma)I, \\
\frac{dB}{dt} &\leq \xi N^* + (g - \delta)B.
\end{align*}
\]

Let \( X = (I, B)^T \). Equation (24) gives

\[ \frac{dX}{dt} \leq (F - V)X. \]

By the Perron–Frobenius theorem, there exists a non-negative left eigenvector \( u \) of the non-negative matrix \( V^{-1}F \) with respect to the eigenvalue \( R_0 = \rho(FV^{-1}) = \rho(V^{-1}F) \). Motivated by [33], we define the Lyapunov function:

\[ L = u^T V^{-1}x. \]

Differentiating \( L \) along solutions of Equation (1) gives

\[ L' = u^T V^{-1} \frac{dX}{dt} \leq u^T V^{-1} (F - V)X = (R_0 - 1)u^T X. \]

If \( R_0 < 1 \), then \( L' \leq 0 \), and \( L' = 0 \) implies \( u^T X = 0 \). Hence either (1) \( I = 0 \) or (2) \( B = 0 \). It follows from the second to the fourth equations of Equation (1) and assumptions \((H1)-(H3)\).
that either (1) \( B = 0 \) or (2) \( I = 0 \). Hence \( I = B = 0 \). In view of the third and the first equations of Equation (1), we see that \( R = 0 \) and \( S = N^* \). Thus, the largest invariant set where \( L' = 0 \) is the singleton \((N^*,0,0,0)\). Therefore, by LaSalle’s Invariant Principle \([16]\), the DFE is globally asymptotically stable in \( \Delta \) if \( R_0 < 1 \).

Case (2): \( R_0 > 1 \).

By Theorem 2.1, the system (1) has two equilibria: the DFE and the EE, and the DFE is unstable. We now employ the geometric approach based on the second compound matrix \([20]\) to analyse the endemic global stability of Equation (1). Essential assumptions and results of the geometric approach are provided in Appendix 2. Since the DFE is unstable and located on the boundary of \( \Delta \), Equation (1) is uniformly persistent; that is, there is a constant \( c > 0 \), such that

\[
\liminf_{t \to \infty} \{ S(t), I(t), R(t), B(t) \} > c.
\]

The compactness of \( \Delta \) and the uniform persistence of the system (1) imply that this system has a compact absorbing set. On the other hand, Equation (1) has a unique equilibrium in the interior of \( \Delta \) in Case (2) by Theorem 2.1. We proceed to verify the Bendixson criterion \( \tilde{q}_2 < 0 \), where \( \tilde{q}_2 \) is as defined in Equation (A8) in Appendix 2.

The Jacobian matrix of the system (1), after dropping the equation for \( R \), at the EE is

\[
J = \begin{bmatrix}
-(d + \sigma) - (f_1(I) + f_2(B)) & -SF'_1(I) - \sigma & -SF'_2(B) \\
(f_1(I) + f_2(B)) & SF'_1(I) - (d + \gamma) & SF'_2(B) \\
0 & \xi & h'(B) - \delta
\end{bmatrix}
\]

(28)

and the associated second compound matrix is

\[
J^{[2]} = \begin{bmatrix}
-(d + \gamma) + \theta(I,B) + \delta + SF'_1(I) & SF'_1(I) & SF'_2(B) \\
\xi & \theta(I,B) + h'(B) & -SF'_1(I) - \sigma \\
0 & (f_1(I) + f_2(B)) & SF'_1(I) + h'(B) - (d + \gamma + \delta)
\end{bmatrix}
\]

(29)

where \( \theta(I,B) = -(d + \sigma + \delta) - (f_1(I) + f_2(B)) \).

Define

\[
P = \text{diag} \left[ 1, \frac{I}{B}, \frac{I}{B} \right].
\]

Then

\[
P_F P^{-1} = \text{diag} \left[ 0, \frac{I}{I - B'}, \frac{I}{I - B'} \right]
\]

and

\[
P J^{[2]} P^{-1} = \begin{bmatrix}
-(d + \gamma) + \theta(I,B) + \delta + SF'_1(I) & \frac{SBF'_1(I)}{I} & \frac{SBF'_2(B)}{I} \\
\frac{I \xi}{B} & \theta(I,B) + h'(B) & -SF'_1(I) - \sigma \\
0 & (f_1(I) + f_2(B)) & SF'_1(I) + h'(B) - (d + \gamma + \delta)
\end{bmatrix}
\]

(30)

The matrix \( Q = P_F P^{-1} + P J^{[2]} P^{-1} \) can be written in the block form as follows:

\[
Q = \begin{bmatrix}
Q_{11} & Q_{12} \\
Q_{21} & Q_{22}
\end{bmatrix}
\]

(31)
in which
\[
Q_{11} = -(d + \gamma) + \theta(I, B) + \delta + Sf'(I),
\]
\[
Q_{12} = \begin{bmatrix}
SBf'_1(B) \\
SBf'_2(B)
\end{bmatrix},
\]
\[
Q_{21} = \begin{bmatrix}
\frac{I^*_1}{B} \\
\frac{I^*_2}{B}
\end{bmatrix},
\]
\[
Q_{22} = \begin{bmatrix}
\theta(I, B) + h'(B) + \frac{I'}{I} - \frac{B'}{B} & -Sf'_1(I) - \sigma \\
f_1(I) + f_2(B) & Sf'_1(I) + h'(B) - (d + \gamma + \delta) + \frac{I'}{I} - \frac{B'}{B}
\end{bmatrix}.
\]

We now define the vector norm for \( \mathbb{R}^3 \) as
\[
|(x_1, x_2, x_3)| = \max\{|x_1|, |x_2| + |x_3|\}
\]
for any \((x_1, x_2, x_3) \in \mathbb{R}^3\). Let \( \mu \) denote the Lozinskii measure with respect to this norm. By direct calculation, we find
\[
\mu(Q) = \sup\{g_1, g_2\}
\]
with
\[
g_1 = \mu_1(Q_{11}) + |Q_{12}|, \\
g_2 = |Q_{21}| + \mu_1(Q_{22}),
\]
where \(|Q_{12}|\) and \(|Q_{21}|\) are the matrix norms induced by \( L_1 \) norm, \( \mu_1 \) denotes the Lozinskii measure with respect to \( L_1 \) norm. Specifically,
\[
g_1 = Q_{11} + |Q_{12}| = -(d + \gamma) + \theta(I, B) + \delta + Sf'_1(I) + \frac{SBf_2(B)}{I},
\]
\[
g_2 = |Q_{21}| - (d + \sigma + \delta) + h'(B) + \frac{I'}{I} - \frac{B'}{B} + \sup\{0, 2[Sf'_1(I) + \sigma] - \gamma\}.
\]

By assumptions (H1) and (H2),
\[
f'_1(I)I \leq f_1(I), \quad f'_2(B)B \leq f_2(B)
\]
for all \((S, I, B)\). Note that \(I'/I = Sf_1(I)/I + Sf_2(B)/I - (d + \gamma)\). It gives
\[
-(d + \gamma) = \frac{I'}{I} - \left[\frac{Sf_1(I)}{I} + \frac{Sf_2(B)}{I}\right].
\]
Thus, we obtain
\[
g_1 = -(d + \sigma) + \frac{I'}{I} - \frac{S}{I}f_1(I) - \frac{S}{I}f_2(B) - (f_1(I) + f_2(B)) + Sf'_1(I) + \frac{SB}{I}f'_2(B),
\]
\[
= \frac{I'}{I} - (d + \sigma) - (f_1(I) + f_2(B)) - \frac{S}{I}((f_2(B) - f'_2(B)B) + (f_1(I) - f'_1(I)I)),
\]
\[
\leq \frac{I'}{I} - (d + \sigma).
\]
(32)
By the similar argument together with assumption (H3), we find

\[ g_2 = \frac{I}{B} \xi - (d + \sigma + \delta) + h'(B) + \frac{I'}{I} - \frac{B'}{B} + \sup\{0, 2(Sf'_I(I) + \sigma) - \gamma\}, \]

\[ \leq \frac{I'}{I} - (d + \sigma) + \sup\{0, 2(Sf'_I(I) + \sigma) - \gamma\}. \]

If \( \sup\{2(Sf'_I(I))\} + \sigma \leq \gamma \), then

\[ \sup\{2(Sf'_I(I) + \sigma)\} - \gamma \leq \sigma. \]

It gives

\[ g_2 \leq \frac{I'}{I} - d. \quad (33) \]

By Equations (32) and (33),

\[ \mu(Q) \leq \frac{I'}{I} - d. \]

In view of \( 0 \leq I(t) \leq N \), we have

\[ \frac{\ln(I(t)) - \ln(I(0))}{t} \leq \frac{d}{2} \]

for \( t \) sufficiently large. Therefore,

\[ \frac{1}{t} \int_0^t \mu(s) \, ds \leq \frac{1}{t} \int_0^t \left( \frac{I'(s)}{I(s)} - d \right) ds = \frac{\ln(I(t)) - \ln(I(0))}{t} - d \leq -\frac{d}{2}, \]

if \( t \) is large enough. This implies \( \bar{q}_2 \leq -d/2 < 0 \). It completes the proof.

Finally, we mention that, realistically, \( \sigma \ll \gamma \) since for cholera, the typical immunity wanning period is 3–5 years [24], whereas the normal recovery period is about 5–7 days [13]. Hence, the inequality in Theorem 2.2(2) basically sets an upper limit for the rate of change of the direct transmission mode to ensure the global stability of the EE.

3. An SIRS-B cholera epidemic model with spatial movement

As mentioned before, spatial heterogeneity plays an important role in disease transmission and the movement of human hosts and dispersal of pathogens may be critical in shaping a cholera epidemic. Thus, in this section, we will extend our SIRS-B model to a PDE system to investigate the spatial dynamics of cholera.

For a simple start, we consider a 1D spatial domain, \( 0 \leq x \leq 1 \), and we assume that both the human population and the bacteria undergo a diffusion process. Let \( D_i > 0 \) (\( 1 \leq i \leq 4 \)) be the diffusion coefficients of \( S, I, R \) and \( B \), respectively. Then the cholera model (1) with inclusion of
diffusion takes the form:

\[
\begin{align*}
\frac{\partial S}{\partial t} &= b - S f_1(I) - S f_2(B) - dS + \sigma R + D_1 \frac{\partial^2 S}{\partial x^2}, \\
\frac{\partial I}{\partial t} &= S f_1(I) + S f_2(B) - (d + \gamma) I + D_2 \frac{\partial^2 I}{\partial x^2}, \\
\frac{\partial R}{\partial t} &= \gamma I - (d + \sigma) R + D_3 \frac{\partial^2 R}{\partial x^2}, \\
\frac{\partial B}{\partial t} &= \xi I + h(B) - \delta B + D_4 \frac{\partial^2 B}{\partial x^2}.
\end{align*}
\]  

(34)

Meanwhile, we assume the entire spatial domain represents a closed community of our interest; that is, no individuals would cross the boundary. Hence, we impose the no-flux Neumann boundary conditions:

\[
\begin{align*}
\frac{\partial S}{\partial x} &= \frac{\partial I}{\partial x} = \frac{\partial R}{\partial x} = \frac{\partial B}{\partial x} = 0, \quad \text{at } x = 0, 1.
\end{align*}
\]  

(35)

Although our spatial domain is oversimplified in some sense, such a PDE modelling framework could be useful in both analytical and applied aspects. For example, the current model, with some further improvement, could be suitable to study the spread of the infection during the first period of the 2010–2012 Haiti cholera outbreak. The suspected source of this outbreak was Artibonite River, the longest as well as the most important river in Haiti, and the initial spread of the disease was along the river [5,41].

### 3.1. Turing instability

One common phenomenon in many reaction–diffusion systems with multiple components is the occurrence of Turing instability [37]; that is, loss of stability due to inclusion of diffusion. In what follows we first study the possibility of a Turing instability regarding our PDE cholera model (34).

We linearize Equation (34) at the DFE, \( p := (y_1^*, y_2^*, y_3^*, y_4^*)^T = (N^*, 0, 0, 0)^T \), of the system (1) in the absence of diffusion. Set

\[
y_i = y_i^* + \delta y_i, \quad (i = 1, 2, 3, 4).
\]

(36)

Let \( Y = (\delta y_1, \delta y_2, \delta y_3, \delta y_4)^T \). The associated linearized system in vector form is given by

\[
\frac{\partial Y}{\partial \tau} = JY + D \frac{\partial^2 Y}{\partial \theta^2},
\]

(37)

where \( J \) is the Jacobian matrix of the associated ODE system evaluated at the DFE, and \( D = \text{diag}[D_1, D_2, D_3, D_4] \).

Consider the eigenvalue problem

\[
\begin{align*}
-\nu_{xx}(x) &= \eta \nu(x), \quad x \in (0, 1), \\
\nu_x(x) &= 0, \quad x = 0, 1.
\end{align*}
\]

(38)
One can easily verify that the eigenvalues \( \eta_k = (k\pi)^2 \geq 0 \) and the corresponding eigenfunctions \( \nu_k(x) = \cos(k\pi x) \). We consider the ansatz
\[
\delta y_i(t, x) = e^{\rho t} \nu(x) \omega_i, \quad (1 \leq i \leq 4),
\]
where \( \nu \) is the solution of the eigenvalue problem (38), and \( \rho \) and \( \omega_i \) are constant. Substituting Equation (39) into Equation (37) yields
\[
(\rho I_4 + \eta D)\omega = J \omega. \quad (40)
\]
We are interested in whether there exists \( \rho \) such that \( \text{Re}(\rho) > 0 \) at the DFE. Solve Equation (40). We find that the associated eigenvalues are given by
\[
\rho_1 = -\frac{(k_1 + k_4)}{2} + \sqrt{\left(4k_2k_3 + \frac{(k_1 - k_4)}{2}\right)^2},
\]
\[
\rho_2 = -\frac{(k_1 + k_4)}{2} - \sqrt{\left(4k_2k_3 + \frac{(k_1 - k_4)}{2}\right)^2},
\]
\[
\rho_3 = -(\eta D_1 + d),
\]
\[
\rho_4 = -(\eta D_3 + d + \sigma),
\]
where \( k_1 = (d + \gamma) - N^*f_1 + \eta D_2, k_2 = N^*f_B, k_3 = \xi \) and \( k_4 = \delta - g + \eta D_4 \).

**Proposition 3.1** If \( R_0 < 1 \), inclusion of diffusive spatial spread into model (1) will not produce a Turing instability.

**Proof** Suppose \( R_0 < 1 \). The ODE model (1) has a unique DFE, and \( (d + \gamma) - N^*f_1 > 0 \) and \( \delta > g \). This gives \( k_i > 0 \) for \( i = 1, 2, 3, 4 \). Thus, \( \rho_2, \rho_3 \) and \( \rho_4 \) are all negative and have the same sign as the corresponding eigenvalue of \( J \). The only eigenvalue that could have a sign change is \( \rho_1 \). Thus, it suffices to study the sign of \( \rho_1 \).

Notice that
\[
\rho_1\rho_2 = k_1k_4 - k_2k_3
= ((d + \gamma) - N^*f_1 + \eta D_2)(\delta - g + \eta D_4) - \xi N^*f_B
= ((d + \gamma) - N^*f_1)D_4 + (\delta - g)D_2 + D_2D_4 + ((d + \gamma) - N^*f_1)(\delta - g) - \xi N^*f_B. \quad (42)
\]
It is clear that in the last equation of Equation (42), the first three terms are positive. By the proof of Theorem 2.1, \( ((d + \gamma) - N^*f_1)(\delta - g) - \xi N^*f_B \) is a product of two eigenvalues at the DFE, and hence it is positive. This yields \( \rho_1\rho_2 > 0 \). Hence, \( \rho_2 < 0 \) implies \( \rho_1 < 0 \). Therefore, all the eigenvalues associated with \( p_0 \) will keep the same sign, and Turing instability will not occur.

Proposition 3.1 states that inclusion of diffusion spatial spread in the original ODE system (1) tends to stabilize the system. An implication is that the PDE model (34) would exhibit threshold dynamics similar to that of the ODE system. The question now is how to quantify the threshold in the spatial–temporal domain for our PDE cholera model. Below we will focus our attention on this point.
3.2. Disease threshold

In contrast to the large body of work devoted to the threshold dynamics of ODE epidemic models, there are very few such studies on PDE models. In a recent article by Wang and Zhao [39], the concept of the basic reproduction number is extended to reaction–diffusion epidemic systems with Neumann (no-flux) boundary conditions. Based on the theory of principle eigenvalues, the authors defined the basic reproduction number $R_0$ for a reaction–diffusion epidemic system as the spectral radius of the operator

$$L[\phi(x)] = \int_0^\infty F(x)T(t)\phi \, dt = F(x) \int_0^\infty T(t)\phi \, dt. \quad (43)$$

Consequently, they showed that

$$\int_0^\infty T(t)\phi \, dt = -B^{-1}\phi, \quad (44)$$

where $B := \nabla \cdot (d_I \nabla) - V$. It then follows that

$$L = -FB^{-1}. \quad (45)$$

Here $F$ and $V$ are analogues to the next-generation matrices associated with the corresponding ODE system (i.e. without diffusion terms), $T(t)$ is the solution semi-group for the linearized reaction–diffusion system, $\phi$ denotes the distribution of the initial infection, and $d_I$ is the diffusion coefficient vector.

For our cholera model (34), we have

$$B = \begin{bmatrix} D_2 \frac{\partial^2}{\partial x^2} - (d + \gamma) & 0 \\ 0 & D_4 \frac{\partial^2}{\partial x^2} - \delta \end{bmatrix}. \quad (46)$$

In order to analyse the basic reproduction number of the PDE system, $R_0^{PDE} = \rho(L)$, we proceed to calculate $B^{-1}$ by solving $B(\phi_1, \phi_2)^T = (y_1, y_2)^T$ subject to homogeneous Neumann boundary conditions. Let us first consider the equation

$$B_1[\phi_1] := D_2 \frac{\partial^2 \phi_1}{\partial x^2} - (d + \gamma)\phi_1 = y_1, \quad 0 \leq x \leq 1;$$

$$\phi_1'(0) = 0, \quad \phi_1'(1) = 0. \quad (47)$$

This problem can be conveniently solved by using the Laplace transform. Denote the Laplace transforms of $\phi_1(x)$ and $y_1(x)$ by $\Phi_1(s)$ and $Y_1(s)$, respectively. We then obtain

$$\Phi_1(s) = \frac{Y_1(s)}{D_2s^2 - (d + \gamma)} + \frac{sD_2\phi_1(0)}{D_2s^2 - (d + \gamma)},$$

where we have applied the first boundary condition of $\phi_1$. The inverse Laplace transform then yields

$$\phi_1(x) = \frac{1}{\sqrt{D_2(d + \gamma)}} \int_0^x \sinh \left[ \sqrt{\frac{d + \gamma}{D_2}} (x - \tau) \right] y_1(\tau) \, d\tau + \phi_1(0) \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} x \right).$$
Now differentiating $\phi_1$ and using the second boundary condition, we obtain

$$\phi_1(0) = \frac{-1}{\sqrt{D_2(d + \gamma) \sinh((d + \gamma)/D_2)}} \int_0^1 \cosh \left[ \sqrt{\frac{d + \gamma}{D_2}} (1 - \tau) \right] y_1(\tau) \, d\tau.$$  

Hence,

$$\phi_1(x) = B_1^{-1}[y_1] = \frac{1}{\sqrt{D_2(d + \gamma)}} \int_0^x \sinh \left[ \sqrt{\frac{d + \gamma}{D_2}} (x - \tau) \right] y_1(\tau) \, d\tau - \frac{\cosh \left( \sqrt{\frac{d + \gamma}{D_2}} x \right)}{\sqrt{D_2(d + \gamma) \sinh((d + \gamma)/D_2)}} \int_0^1 \cosh \left[ \sqrt{\frac{d + \gamma}{D_2}} (1 - \tau) \right] y_1(\tau) \, d\tau. \quad (48)$$

In this expression, the second part represents the homogeneous solution, whereas the first part represents a particular solution to the original non-homogeneous equation. In a similar way, we can solve the boundary value problem

$$B_2[\phi_2] := D_4 \frac{\partial^2 \phi_2}{\partial x^2} - \delta \phi_2 = y_2, \quad 0 \leq x \leq 1;$$

$$\phi'_2(0) = 0, \quad \phi'_2(1) = 0 \quad (49)$$

to obtain

$$\phi_2(x) = B_2^{-1}[y_2] = \frac{1}{\sqrt{D_4 \delta}} \int_0^x \sinh \left[ \sqrt{\frac{\delta}{D_4}} (x - \tau) \right] y_2(\tau) \, d\tau - \frac{\cosh \left( \sqrt{\frac{\delta}{D_4}} x \right)}{\sqrt{D_4 \delta \sinh((\delta)/D_4)}} \int_0^1 \cosh \left[ \sqrt{\frac{\delta}{D_4}} (1 - \tau) \right] y_2(\tau) \, d\tau. \quad (50)$$

For consistence in notations, below we will switch $\phi_1$ and $y_1$ in Equation (48), and $\phi_2$ and $y_2$ in Equation (50). Now we look at the eigenvalue problem $L[\phi] = \lambda \phi$; that is,

$$-FB^{-1} \phi = \lambda \phi, \quad (51)$$

where, in our cholera model,

$$F = \begin{bmatrix} N^*f_1 & N^*f_2 \\ \xi & g \end{bmatrix} = \begin{bmatrix} N^*f_1(0) & N^*f_2(0) \\ \xi & h'(0) \end{bmatrix}. \quad (52)$$

We thus obtain the following two equations

$$k_{i1} \int_0^x \sinh \left[ \sqrt{\frac{d + \gamma}{D_2}} (x - \tau) \right] \phi_1(\tau) \, d\tau + k_{i2} \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} x \right) \int_0^1 \cosh \left[ \sqrt{\frac{d + \gamma}{D_2}} (1 - \tau) \right] \phi_1(\tau) \, d\tau$$

$$+ k_{i3} \int_0^x \sinh \left[ \sqrt{\frac{\delta}{D_4}} (x - \tau) \right] \phi_2(\tau) \, d\tau + k_{i4} \cosh \left( \sqrt{\frac{\delta}{D_4}} x \right) \int_0^1 \cosh \left[ \sqrt{\frac{\delta}{D_4}} (1 - \tau) \right] \phi_2(\tau) \, d\tau = \lambda \phi_i(x), \quad (i = 1, 2). \quad (53)$$
with the coefficients

\[

k_{11} = -\frac{N^* f_I}{\sqrt{D_2 (d + \gamma)}}, \quad k_{12} = \frac{N^* f_I}{\sqrt{D_2 (d + \gamma) \sinh(\sqrt{d + \gamma}/D_2)}},
\]

\[

k_{13} = -\frac{N^* f_I}{\sqrt{D_4 \delta}}, \quad k_{14} = \frac{N^* f_I}{\sqrt{D_4 \delta \sinh(\sqrt{\delta}/D_4)}},
\]

\[

k_{21} = -\frac{\xi}{\sqrt{D_2 (d + \gamma)}}, \quad k_{22} = \frac{\xi}{\sqrt{D_2 (d + \gamma) \sinh(\sqrt{d + \gamma}/D_2)}},
\]

\[

k_{23} = -\frac{g}{\sqrt{D_4 \delta}}, \quad k_{24} = \frac{g}{\sqrt{D_4 \delta \sinh(\sqrt{\delta}/D_4)}}.
\]

Analysis of the eigenvalue problem associated with the integral equations (53) appears difficult. Instead, we may gain some insights into the eigenvalue problem (53) by looking at its discrete form. Let us partition the interval \([0, 1]\) uniformly with the nodes \(x_i = i \Delta x (0 \leq i \leq M)\), where \(M/\Delta x = 1\). Evaluating Equations (53) at \(x_i\) for \(i = 1, 2, \ldots, M\), and simply approximating each integral by using the right endpoint of each subinterval, we obtain a matrix equation

\[
AY = \lambda Y \tag{54}
\]

with \(Y = [\phi_1(x_1), \phi_1(x_2), \ldots, \phi_1(x_M), \phi_2(x_1), \phi_2(x_2), \ldots, \phi_2(x_M)]^T\). The problem is now reduced to analysing the spectral radius, \(\rho(A)\), of the coefficient matrix \(A\); in particular, we have \(\lim_{\Delta x \to 0} \rho(A) = \rho(L) = R_{0}^{\text{PDE}}\).

The coefficient matrix \(A\) in Equation (54) can be written as

\[
A = A_1 + A_2 + A_3 + A_4, \tag{55}
\]

where each matrix \(A_i\) (1 \(\leq i \leq 4\), of dimension \(2M \times 2M\), results from the discretization of the \(i\)th integral in Equations (53). Specifically, \(A_1\) can be represented by a block form

\[
A_1 = \Delta x \begin{bmatrix} k_{11} \tilde{A}_1 & 0 \times M \\ k_{21} \tilde{A}_1 & 0 \times M \end{bmatrix}, \tag{56}
\]

where \(0 \times M\) denotes the zero square matrix of dimension \(M \times M\), and \(\tilde{A}_1 = (a_{1,ij})\) is an \(M \times M\) lower-triangular matrix given by

\[
a_{1,ij} = \begin{cases} \sinh \left[ \sqrt{\frac{d + \gamma}{D_2}} (x_i - x_j) \right] & \text{if } i \geq j, \\ 0, & \text{otherwise}. \end{cases} \tag{57}
\]

Similarly, we have

\[
A_3 = \Delta x \begin{bmatrix} 0 \times M & k_{13} \tilde{A}_3 \\ 0 \times M & k_{23} \tilde{A}_3 \end{bmatrix}, \tag{58}
\]

where \(\tilde{A}_3 = (a_{3,ij})\) is also an \(M \times M\) lower-triangular matrix:

\[
a_{3,ij} = \begin{cases} \sinh \left[ \sqrt{\frac{\delta}{D_4}} (x_i - x_j) \right] & \text{if } i \geq j, \\ 0, & \text{otherwise}. \end{cases} \tag{59}
\]

The matrix \(A_2\) takes the form

\[
A_2 = \Delta x \begin{bmatrix} k_{12} \tilde{A}_2 & 0 \times M \\ k_{22} \tilde{A}_2 & 0 \times M \end{bmatrix}. \tag{60}
\]
with the $M \times M$ block $\tilde{A}_2 = (a_{2,ij})$ for which
\[
a_{2,ij} = \cosh\left(\sqrt{\frac{d + \gamma}{D_2}} x_i\right) \cosh\left(\sqrt{\frac{d + \gamma}{D_2}} (1 - x_j)\right), \quad 1 \leq i, j \leq M.
\]

Finally, The matrix $A_4$ takes the form
\[
A_4 = \Delta x \begin{bmatrix} 0_M & k_{14} \tilde{A}_4 \\ 0_M & k_{24} \tilde{A}_4 \end{bmatrix}
\]
with the $M \times M$ block $\tilde{A}_4 = (a_{4,ij})$ for which
\[
a_{4,ij} = \cosh\left(\sqrt{\frac{\delta}{D_4}} x_i\right) \cosh\left(\sqrt{\frac{\delta}{D_4}} (1 - x_j)\right), \quad 1 \leq i, j \leq M.
\]

It is obvious that the spectral radius for each of $A_1$ and $A_3$ is 0; that is,
\[
\rho(A_1) = \rho(A_3) = 0. \tag{62}
\]

Meanwhile, using Lemma A.1 from Appendix 3, we obtain the characteristic polynomial of $\tilde{A}_2$ as
\[
\det(\lambda I_M - \tilde{A}_2) = \lambda^{M-1}(\lambda - \tilde{\lambda}_2), \tag{63}
\]
where
\[
\tilde{\lambda}_2 := \sum_{i=1}^M \cosh\left(\sqrt{\frac{d + \gamma}{D_2}} x_i\right) \cosh\left(\sqrt{\frac{d + \gamma}{D_2}} (1 - x_i)\right). \tag{64}
\]

Note that $\tilde{A}_2$ is a strictly positive matrix and $\tilde{\lambda}_2$ is the unique positive dominant eigenvalue of $\tilde{A}_2$. Similarly, the matrix $\tilde{A}_4$ is also strictly positive and we have
\[
\det(\lambda I_M - \tilde{A}_4) = \lambda^{M-1}(\lambda - \tilde{\lambda}_4), \tag{65}
\]
where
\[
\tilde{\lambda}_4 := \sum_{i=1}^M \cosh\left(\sqrt{\frac{\delta}{D_4}} x_i\right) \cosh\left(\sqrt{\frac{\delta}{D_4}} (1 - x_i)\right) \tag{66}
\]
is the unique positive dominant eigenvalue of $\tilde{A}_4$. It is clear to see that
\[
\rho(A_2) = \Delta x k_{12} \tilde{\lambda}_2, \quad \rho(A_4) = \Delta x k_{24} \tilde{\lambda}_4. \tag{67}
\]

With some algebraic manipulation, we can observe that when $\Delta x \to 0$,
\[
\lim_{\Delta x \to 0} \rho(A_2) = \frac{k_{12}}{2} \cosh\left(\sqrt{\frac{d + \gamma}{D_2}}\right) + \frac{k_{12}}{2} \sqrt{\frac{D_2}{d + \gamma}} \sinh\left(\sqrt{\frac{d + \gamma}{D_2}}\right) \approx \frac{N f_I}{d + \gamma}, \tag{68}
\]
provided that $(d + \gamma)/D_2 \ll 1$. In a similar way,
\[
\lim_{\Delta x \to 0} \rho(A_4) = \frac{k_{24}}{2} \cosh\left(\sqrt{\frac{\delta}{D_4}}\right) + \frac{k_{24}}{2} \sqrt{\frac{D_4}{\delta}} \sinh\left(\sqrt{\frac{\delta}{D_4}}\right) \approx \frac{g}{\delta}, \tag{69}
\]
provided that $\delta/D_4 \ll 1$. Note that $N f_I/(d + \gamma)$ and $g/\delta$ are two components in the expression of $R_0$ for our ODE model; indeed, in the ODE model, we have $R_0 \geq \max(N f_I/(d + \gamma), g/\delta)$. 


There is no general relationship between $\rho(A)$ and $\rho(A_i)$ ($1 \leq i \leq 4$). Nevertheless, if we assume $(d + \gamma)/D_2 \ll 1$ and $\delta/D_4 \ll 1$, then each entry of $A_1$ and $\tilde{A}_3$ is very small; it is bounded between 0 and $\sinh((d + \gamma)/D_2^2)$ if in $A_1$, and between 0 and $\sinh(\delta/D_4)$ if in $A_3$. Thus, $A_1$ and $A_3$ can be treated as small perturbation to $A_2$ and $A_4$ in this case, and

$$\rho(A) \approx \rho(A_2 + A_4).$$

(70)

Next we explore lower and upper bounds for $\rho(A_2 + A_4)$. We first have (see, e.g. [15])

$$\max(\rho(A_2), \rho(A_4)) \leq \rho(A_2 + A_4),$$

(71)

by noting that both $A_2$ and $A_4$ are non-negative matrices. In addition, some standard results on the lower/upper bounds of non-negative matrices based on minimum/maximum row (or, column) sums [15] can be applied to estimate $\rho(A_2 + A_4)$ and give a range for $\rho(L)$. In order to compare the ODE and PDE thresholds, however, we will make use of the following result [42] to determine an upper bound for $\rho(A_2 + A_4)$.

**Lemma 3.1** Consider a non-negative irreducible square matrix $Q$ of the block form $Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix}$, where $Q_1$ and $Q_4$ are both square matrices. If $\rho(Q) > \|Q_1\|$, where $\| \cdot \|$ denotes any consistent matrix norm, then

$$\rho(Q) \leq \frac{1}{2}((\|Q_1\| + \|Q_4\|) + \sqrt{(\|Q_1\| - \|Q_4\|)^2 + 4\|Q_2\|\|Q_3\|}).$$

(72)

If we treat the matrix $A_2 + A_4$ as $Q$ in the lemma above, then we have $Q_1 = \Delta x k_{12} \tilde{A}_2$, $Q_2 = \Delta x k_{14} \tilde{A}_4$, $Q_3 = \Delta x k_{22} \tilde{A}_2$, $Q_4 = \Delta x k_{24} \tilde{A}_4$. Since $A_2 + A_4$ is strictly positive, $Q$ must be irreducible. For simplicity, let us use the induced $\| \cdot \|_1$ norm; that is, the maximum column sum of each (non-negative) matrix. Through direct calculation, it is easy to obtain the following:

$$\lim_{\Delta x \to 0} \|Q_1\|_1 = k_{12} \sqrt{\frac{D_2}{d + \gamma}} \sinh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) = \frac{Nf_B}{d + \gamma} \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) := P_1,$$

$$\lim_{\Delta x \to 0} \|Q_3\|_1 = k_{22} \sqrt{\frac{D_2}{d + \gamma}} \sinh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) = \frac{\xi}{d + \gamma} \cosh \left( \sqrt{\frac{d + \gamma}{D_2}} \right) := P_3,$$

$$\lim_{\Delta x \to 0} \|Q_2\|_1 = k_{14} \sqrt{\frac{D_4}{\delta}} \sinh \left( \sqrt{\frac{\delta}{D_4}} \right) \cosh \left( \sqrt{\frac{\delta}{D_4}} \right) = \frac{Nf_B}{\delta} \cosh \left( \sqrt{\frac{\delta}{D_4}} \right) := P_2,$$

$$\lim_{\Delta x \to 0} \|Q_4\|_1 = k_{24} \sqrt{\frac{D_4}{\delta}} \sinh \left( \sqrt{\frac{\delta}{D_4}} \right) \cosh \left( \sqrt{\frac{\delta}{D_4}} \right) = \frac{g}{\delta} \cosh \left( \sqrt{\frac{\delta}{D_4}} \right) := P_4.$$

Meanwhile, the condition $\rho(Q) > \|Q_1\|_1$ is satisfied based on Equations (68) and (71). We thus obtain, as $\Delta x \to 0$,

$$\rho(A_2 + A_4) \leq \frac{1}{2}[P_1 + P_4 + \sqrt{(P_1 - P_4)^2 + 4P_2P_3}].$$

(73)
Furthermore, if we let \((d + \gamma) / D_2 \to 0\) and \(\delta / D_4 \to 0\), we can easily observe that

\[
\rho(A_2 + A_4) \leq \frac{1}{2} \left[ \frac{Nf_I}{d + \gamma} + \frac{g}{\delta} + \sqrt{\left( \frac{Nf_I}{d + \gamma} - \frac{g}{\delta} \right)^2 + 4 \frac{\xi Nf_B}{\delta (d + \gamma)}} \right] := R_{0}^{\text{ODE}}. \tag{74}
\]

Note that the right-hand side of the inequality (74) gives \(R_0\) for our ODE model. The results in Equations (73) and (74) imply that when the diffusion coefficients \(D_2\) and \(D_4\) are not so large, the value of \(\rho(L)\) can possibly exceed \(R_{0}^{\text{ODE}}\), since \(\cosh(x) > 1\) for \(x > 0\). However, when \(D_2\) and \(D_4\) are very large, we have \(R_{0}^{\text{PDE}} = \rho(L) \leq R_{0}^{\text{ODE}}\); that is, the basic reproduction number for our PDE model would be equal to or lower than that for the ODE model. Indeed, we can establish a even stronger result regarding the limit case.

**Proposition 3.2** When \((d + \gamma) / D_2 \to 0\) and \(\delta / D_4 \to 0\), we have

\[
R_{0}^{\text{PDE}} = R_{0}^{\text{ODE}}.
\]

**Proof** Within such limits, the eigenvalue problem (53) is reduced to

\[
\frac{Nf_I}{d + \gamma} \int_0^1 \phi_1(\tau) \, d\tau + \frac{Nf_B}{\delta} \int_0^1 \phi_2(\tau) \, d\tau = \lambda \phi_1(x),
\]

\[
\frac{\xi}{d + \gamma} \int_0^1 \phi_1(\tau) \, d\tau + \frac{g}{\delta} \int_0^1 \phi_2(\tau) \, d\tau = \lambda \phi_2(x).
\]

From these two equations, it is straightforward to see that \(\phi_1(x) \equiv c_1\) and \(\phi_2(x) \equiv c_2\), where \(c_1\) and \(c_2\) are some nonzero constants. Let \(Z = c_2 / c_1\). Then we have

\[
\lambda = \frac{Nf_I}{d + \gamma} + \frac{Nf_B}{\delta} Z = \frac{g}{\delta} + \frac{\xi}{d + \gamma} \frac{1}{Z}.
\]

Solving the quadratic equation for \(Z\), we obtain

\[
Z_{\pm} = \frac{-(Nf_I/(d + \gamma) - g/\delta) \pm \sqrt{(Nf_I/(d + \gamma) - g/\delta)^2 + 4(\xi Nf_B/\delta (d + \gamma))}}{2(Nf_B/\delta)}.
\]

Consequently, we have

\[
\lambda_{\pm} = \frac{1}{2} \left[ \frac{Nf_I}{d + \gamma} + \frac{g}{\delta} \pm \sqrt{\left( \frac{Nf_I}{d + \gamma} - \frac{g}{\delta} \right)^2 + 4 \frac{\xi Nf_B}{\delta (d + \gamma)}} \right].
\]

Clearly, \(|\lambda_+| > |\lambda_-|\) and \(\rho(L) = \lambda_+ > 0\). We thus obtain \(R_{0}^{\text{PDE}} = \lambda_+ = R_{0}^{\text{ODE}}. \)

As one can naturally expect, when the diffusion terms are incorporated into the original ODE model (i.e. spatial variation and human/pathogen movement are taken into account), the infection risk predicted by the model will be changed, leading to a different basic reproduction number. The result in Proposition 3.2, nevertheless, implies that when the diffusion is extremely strong, it tends to smooth out the spatial heterogeneity and reduce the system to a homogeneous setting, such that the ODE and PDE models would share the same disease threshold.
4. Numerical results

We now verify our analytical results by numerical simulation. In particular, we will numerically demonstrate (and compare) the threshold dynamics based on our ODE and PDE cholera models. To conduct the simulation, the direct and indirect transmissions are assumed to take the form developed in Mukandavire et al. [22]:

\[ f_1(B) = \beta_h I \] and \[ f_2(B) = \beta_e B/(B + \kappa) \], where \( \beta_h \) and \( \beta_e \) represent the direct and indirect transmission parameters, respectively; \( \kappa \) is the concentration of \textit{V. cholera} in the environment depicting the half saturation rate. Cholera transmission through ingestion of bacteria from a contaminated environment can be greatly influenced by growth and survival of free-living bacteria in the environment [27]. Therefore, growth of the pathogen in the environment cannot be ignored in a realistic model. We include bacteria growth through a logistic growth assumption [23], that is,

\[ h(B) = gB \left( 1 - \frac{B}{K} \right) \],

where \( g = h'(0) \) is the intrinsic bacteria growth rate and \( K \) is the maximal capacity of free-living bacteria in the environment.

The model parameters of Equation (1) are listed in Table A1. Our particular interest is the impact of the bacterial growth rate, \( g \), on the value of \( R_0 \). Meanwhile, the shedding rate, \( \xi \), describes how much an infected individual contributes to the bacterial concentration in the water reservoir, and is likely to vary widely [12]. Thus, we plot \( R_0 \) associated with the ODE model (1) as a function of \( g \) and \( \xi \) in Figure 2(a). Particularly, the black curve, denoted as \( \Gamma \),

\[ \xi = w(g) := \frac{(\delta - g)(d + \gamma - N^*f_I)}{N^*f_B} \],

indicates where \( R_0 = 1 \). As one can see from Figure 2(a) the bacterial growth rate \( g \) and shedding rate \( \xi \) play a vital role in shaping \( R_0 \). In particular, (i) When \( g < \delta \), the disease threshold changes from \( R_0 < 1 \) to \( R_0 > 1 \) provided that either of the following happens:

1. \( g \) is fixed at a constant value in the range and \( \xi \) increases to cross \( \Gamma \);

2. \( \xi < w(0) = 10.85 \text{cells} \cdot \text{ml}^{-1} \cdot \text{week}^{-1} \) is fixed and \( g \) increases to cross \( \Gamma \).

(ii) When \( g \geq \delta = \frac{7}{30} \text{week}^{-1} \), \( R_0 > 1 \). Moreover, for each fixed value of \( g \), the bacteria shedding rate \( \xi \) becomes less important and even irrelevant as \( g \) increases, and for each fixed value of \( \xi \), varying \( g \) can lead to a significant change in the value of \( R_0 \). An illustration of the number of
Figure 3. The difference in the basic reproduction number between the PDE model (34) and the ODE model (1): \( R_0^{\text{PDE}} - R_0^{\text{ODE}} \).

Figure 4. Number of infected hosts associated with the PDE model (34) vs. space and time. (a) \( R_0 < 1 \); (b) \( R_0 > 1 \).

infectious individuals as a function of time for \( R_0 < 1 \) and \( R_0 > 1 \) is displayed in Figure 2(b). It clearly shows that the infection dies out when \( R_0 < 1 \), and persists (approaching the EE over time) when \( R_0 > 1 \).

Now let us consider the PDE model (34). Figure 3 shows the difference in the basic reproduction numbers of the PDE model (34) and the ODE model (1). We observe, from Figure 3(a), that \( R_0^{\text{PDE}} - R_0^{\text{ODE}} \) is nearly invisible when the diffusion coefficients of the hosts and bacteria, \( D_2 \) and \( D_4 \), are getting large. This result confirms that \( \lim_{D_2, D_4 \to \infty} R_0^{\text{PDE}} = R_0^{\text{ODE}} \). When \( D_2 \) and \( D_4 \) decrease to zero, Figure 3(b) illustrates that \( R_0^{\text{PDE}} - R_0^{\text{ODE}} > 0 \), and for the fixed bacteria growth rate \( g \) and shedding rate \( \xi \), \( R_0^{\text{PDE}} \) increases and becomes uniformly greater than one for all the possible values of \( g > 0 \) and \( \xi > 0 \). In particular, (i) for a fixed \( g \), \( R_0^{\text{PDE}} - R_0^{\text{ODE}} \) will be elevated if \( \xi \) increases; (ii) for a fixed \( \xi \), \( R_0^{\text{PDE}} - R_0^{\text{ODE}} \) goes up if \( g \) gets larger.

Figure 4 illustrates the number of infectious individuals based on the PDE model (34) with uniform initial distribution, as a function of space and time when the associated \( R_0 \) is lower or higher than the unity. We have also tested other initial conditions with various distribution types, and they all lead to similar patterns in terms of the extinction (when \( R_0 < 1 \)) and endemic state (when \( R_0 > 1 \)) of cholera infection.
5. Discussion

In this paper, we have proposed novel ODE and PDE cholera models that address the intrinsic bacterial dynamics and spatial movement of human and pathogen. Our ODE model employs general incidence functions for both the direct and indirect transmission routes, and incorporates a general representation of the bacterial growth. The resulting basic reproduction number, $R_0$, depends on all these factors. Particularly, we find that bacterial growth contributes to the increase of the value of $R_0$, indicating that the risk of cholera outbreak would be higher than most of prior model predictions in the absence of intrinsic bacterial dynamics. Our model also takes into account the waning cholera immunity in reality, by using an SIRS framework for the human population dynamics. Our analysis shows that, despite the incorporation of bacterial growth and partial immunity, the basic reproduction number remains a sharp threshold for cholera dynamics: when $R_0 < 1$, the DFE is globally asymptotically stable and the disease completely dies out; when $R_0 > 1$, the unique EE is globally asymptotically stable and the disease persists.

Meanwhile, we have extended this ODE model to a PDE system so as to explicitly track the spatial variation on cholera transmission. Our focus again is on the threshold dynamics, and we have formulated the basic reproduction number $R_0$ of this PDE model based on recent work in [39]. We have analysed $R_0$ and estimated its bounds by reducing the operator eigenvalue problem to a matrix eigenvalue problem, which also provides a practical means to compute $R_0$ for the PDE system. As a result of our analysis and simulation, we find that only if the diffusion coefficients are very large, the basic reproduction number of the PDE model, $R_{PDE}^0$, is matching that of the corresponding ODE model, $R_{ODE}^0$; otherwise, $R_{PDE}^0$ is higher than $R_{ODE}^0$. The implication is that, when we take into account the realistic spatial movement of human hosts and pathogen, the overall infection risk might be higher than what would be based on the (unrealistic) homogeneous setting, as the diffusion process may help spread the disease more rapidly. In this regard, using the ODE model might underestimate the epidemic risk. Our findings could provide useful guideline for public health administration to properly scale their efforts in cholera control.

As mentioned before, Joh et al. [18] and Bani-Yaghoub et al. [3] have also considered intrinsic bacterial dynamics in their cholera studies based on ODE systems. The model in [18] considered only the indirect transmission route and focused on logistic growth of the bacteria, whereas the model in [3] investigated the dual transmission pathways with bilinear incidences and logistic bacterial growth. Compared to these, our SIRS-B ODE model is much more general that allows various formulation of the incidence factors and the bacterial growth dynamics. Our results also generalize the findings in these previous studies.

In the present work, we have chosen to model the spatial effects of cholera transmission using a reaction–diffusion PDE system. Similar PDE modelling approach was also used by Bertuzzo et al. [4] in studying the spatial dynamics of cholera. Our PDE model formulation is more general than that of [4] in terms of variety of incidence and pathogen functions, inclusion of multiple transmission pathways, and incorporation of bacterial intrinsic growth. The scope of our work is also different from that in [4] as our focus is on the analysis and simulation of the threshold dynamics for cholera transmission. Another common approach to investigate disease spatial dynamics is based on metapopulation ODE models. Tuite et al. [36] performed a study on the Haiti cholera outbreak using a patch model. In a recent work by Shuai and van den Driessche [31], a multigroup cholera model was presented and analysed. These metapopulation models can potentially reveal more details on the disease dynamics of an individual group (e.g. a province, or a city). The cost, however, is that a large number of groups may be needed for modelling at different spatial scales, which easily leads to a large dynamical system. Consequently, this would pose a challenge for both the analysis and computation of the model. Compared to the cholera models in [31,36], our PDE system appears simpler and is easy to construct, and straightforward for computation. Reducing the operator eigenvalue problem to a matrix eigenvalue problem also allows us to easily calculate the basic reproduction number for our PDE model.
Meanwhile, there are several limitations in our PDE model, and the current study based on our PDE framework can be extended in a number of ways. First, spread of free-living bacteria may occur through the movement of human host, thus a cross-diffusion term would be appropriate in the equation for $B$ in system (34). Practically, infected individuals might not be as active as the susceptible and recovered people, and the value of $D_2$ might be much less than those of $D_1$ and $D_3$. (We note, though, that most cholera infections are minor and asymptomatic, and those people with minor infections could still pursue normal activities including travel and other movements.) Also, spread of free-living bacteria does not necessarily occur on the same spatial scale, which demands more careful modelling and analysis. Moreover, instead of using a simplistic 1D space dimension, the system can be similarly constructed on a 2D spatial domain for more realistic cholera modelling. The diffusion coefficients as well as several parameters of disease transmission rates can be taken as space dependent, instead of constants, to better reflect the details of spatial heterogeneity. In addition, travelling wave solutions can be analysed which will add useful insight into the spatial spread of cholera.

Although we have considered several sources of variability, demographic and environmental, there are other variability factors that could be included in a more comprehensive model. An important example among these is the individual heterogeneity in shedding density and the possibility of super-shedding [30] and hyperinfectivity [13]. The impact of such individual heterogeneity in shedding and infectivity on cholera epidemics will provide an interesting topic in future research.

Acknowledgment

The authors thank the editor and two anonymous referees for their valuable comments that improved this paper.

Funding

This work was partially supported by a grant from the Simons Foundation [#317047 to Xueying Wang]. Jin Wang was partially supported by the National Science Foundation [Grant nos. 1216936, 1245769 and 1412826].

References

[1] A. Alam, R.C. LaRocque, J.B. Harris, C. Vanderspurt, E.T. Ryan, F. Qadri, and S.B. Calderwood, Hyperinfectivity of human-passaged Vibrio cholerae can be modelled by growth in the infant mouse, Infect. Immun. 73 (2005), pp. 6674–6679.
[2] J.R. Andrews and S. Basu, Transmission dynamics and control of cholera in Haiti: An epidemic model, Lancet 377 (2011), pp. 1248–1255.
[3] M. Bani-Yaghoub, R. Gautam, Z. Shuai, P. van den Driessche, and R. Ivanek, Reproduction numbers for infections with free-living pathogens growing in the environment, J. Biol. Dyn. 6 (2012), pp. 923–940.
[4] E. Bertuzzo, R. Casagrandi, M. Gatto, I. Rodriguez-Iturbe, and A. Rinaldo, On spatially explicit models of cholera epidemics, J. Roy. Soc. Interface 7 (2010), pp. 321–333.
[5] Center for Disease Control and Prevention (CDC) web page on Haiti cholera outbreak. Available at: http://www.bt.cdc.gov/situationawareness/haiticholera/dynamicmaps.asp.
[6] D.L. Chao, M.E. Halloran, and I.M. Longini Jr., Vaccination strategies for epidemic cholera in Haiti with implications for the developing world, Proc. Natl. Acad. Sci. 108 (2011), p. 7081–7085.
[7] C.T. Codeço, Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir, BMC Infect. Dis. 1(1) (2001). Available at: http://link.springer.com/article/10.1186/1471-2334-1-1
[8] R.R. Colwell, A global and historical perspective of the genus Vibrio, in The Biology of Vibrios, F.L. Thompson, B. Austin, and J. Swings, eds., ASM Press, Washington, DC, 2006, pp. 3–11.
[9] O. Diekmann, J.A.P. Heesterbeek, and J.A.J. Metz, On the definition and the computation of the basic reproduction ratio $R_0$ in models for infectious diseases in heterogeneous population, J. Math. Biol. 28 (1990), pp. 365–382.
[10] P. van den Driessche and J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (2002), pp. 29–48.
[11] M. Ghosh, P. Chandra, P. Sinha, and J.B. Shukla, Modeling the spread of carrier-dependent infectious diseases with environmental effect, Appl. Math. Comput. 152 (2004), pp. 385–402.
[12] Y.H. Grad, J.C. Miller, and M. Lipsitch. Cholera modeling: Challenges to quantitative analysis and predicting the impacts of interventions, Epidemiology 23 (2012), pp. 523–530.
Appendix 1. Other kinds of reproduction numbers

Type reproduction number. If disease control is targeted at a particular host type, a useful threshold is known as the type reproduction number, \( T \). The type reproduction number defines the expected number of secondary infective cases of a particular population type caused by a typical primary case in a completely susceptible population [14,29]. It is an extension of the basic reproduction number \( R_0 \).
The type reproduction number $T_1$ (resp. $T_2$) for control of infection among humans (resp. bacteria in the contaminated environment) is defined in references [14,29] as

$$T_i = e_i^T K (I - (I - P_i) K)^{-1} e_i,$$  \hspace{1cm} (A1)

provided the spectral radius of matrix $(I - P_i) K$ is less than one, that is, $\rho((I - P_i) K) < 1$, for $i = 1, 2$. Here $I$ is the $2 \times 2$ identity matrix, vector $e_i = (1, 0)^T$, $e_2 = (0, 1)^T$, $K$ is the next-generation matrix, and $P_i$ is the $2 \times 2$ projection matrix with all zero entries except that the $(i, i)$ entry is $1$ $(i = 1, 2)$. Write $K = (k_{ij})$. The type reproduction $T$ can be easily defined in terms of the elements $k_{ij}$:

$$T_1 = k_{11} + \frac{k_{12}k_{21}}{1 - k_{22}},$$

$$T_2 = k_{22} + \frac{k_{12}k_{21}}{1 - k_{11}},$$  \hspace{1cm} (A2)

which exists provided $k_{22} < 1$. Based on Equation (A2), the type reproduction number associated with the infectious humans is

$$T_1 = \frac{N^* f_1}{\delta + \gamma} + \frac{\xi N^* f_B}{(\delta - g)(\delta + \gamma)},$$  \hspace{1cm} (A3)

provided $g < \delta$. The type reproduction number associated with bacteria in the contaminated environment is

$$T_2 = \frac{g}{\delta} + \frac{\xi N^* f_B}{\delta(d + \gamma - N^* f_1)},$$  \hspace{1cm} (A4)

provided $N^* f_1 < d + \gamma$. It is shown in [29] that $R_0 < 1 \equiv T_1 < 1 \equiv T_2 < 1$.

**Target reproduction number.** Another threshold named the target reproduction number, $T_S$, has been recently introduced by Shuai et al. [33]. It can be regarded as an extension of the type reproduction number. Let $S = \{(i_1, j_1), \ldots, (i_m, j_m)\}$ denote the set of the target indices. Define $S_1 = \{i_1, \ldots, i_m\}$ and $S_2 = \{j_1, \ldots, j_m\}$. The target reproduction number is defined as

$$T_S = \rho(E_{S_1} P_{S_1} K P_{S_2} (I - (K - P_{S_1} K P_{S_2}))^{-1} E_{S_1}),$$

provided $\rho(K - P_{S_1} K P_{S_2}) < 1$. Here $E_{S_1} = (e_i)$ is a $2 \times 2$ matrix with $e_i = 1$ if $i = j$ and $i \in S_1$, and zero otherwise. $P_{S_1} = (p_{i,j}^k)$ is a $2 \times 2$ matrix with $p_{i,j}^k = 1$ if $i = j$ and $i \in S_1$, and zero otherwise, for $k = 1, 2$.

There are a number of ways to control the cholera disease. If we target at the vaccination of humans, the target set will be $S = \{(1, 1), (1, 2)\}$ and the associated target reproduction number is the same as the type reproduction number of the infectious hosts, that is,

$$T_S = T_1,$$

provided $g < \delta$. If we target at hygienic disposal of human faeces, then $S = \{(2, 1)\}$ and

$$T_S = T_{21} = \frac{k_{12}k_{21}}{(1 - k_{11})(1 - k_{22})} = \frac{\xi N^* f_B}{(\delta - g)(\delta + \gamma - N^* f_1)},$$

provided $N^* f_1 < d + \gamma$ and $g < \delta$. If we target at the control of the contaminated environment, such as food and water, then $S = \{(1, 2)\}$ and

$$T_S = T_{12} = T_{21},$$

provided $N^* f_1 < d + \gamma$ and $h_E < \delta$. If the isolation of infectious hosts can be done to reduce the direct transmission between humans, then $S = \{(1, 1)\}$ and

$$T_S = T_{11} = \frac{k_{11}(1 - k_{22})}{(1 - k_{22} - k_{12}k_{21})} = \frac{N^* f_1(\delta - g)}{(\delta - g)(\delta + \gamma) - \xi N^* f_B},$$

provided $g(d + \gamma) + \xi N^* f_B < \delta(d + \gamma)$.
Appendix 2. The geometric approach

Here we present the main result of the geometric approach for global stability, originally developed by Li and Muldowney [20].

We consider a dynamical system

$$\frac{dX}{dt} = F(X),$$

(A5)

where $F : D \mapsto \mathbb{R}^n$ is a $C^1$ function and $D \subset \mathbb{R}^n$ is a simply connected open set. Let $P(X)$ be a $\binom{n}{2} \times \binom{n}{2}$ matrix-valued $C^1$ function in $D$, and set

$$Q = PF P^{-1} + PJ^{[2]}P^{-1},$$

(A6)

where $P_F$ is the derivative of $P$ (entry-wise) along the direction of $F$, and $J^{[2]}$ is the second additive compound matrix of the Jacobian $J(X) = DF(X)$. Let $m(Q)$ be the Lozinski\' measure of $Q$ with respect to a matrix norm; that is,

$$m(Q) = \lim_{h \to 0^+} \frac{\|I + hQ\| - 1}{h},$$

(A7)

where $I$ represent the identity matrix. Define a quantity $\bar{q}_2$ as

$$\bar{q}_2 = \limsup_{t \to \infty} \sup_{X_0 \in K} \frac{1}{t} \int_0^t m(Q(X(s,X_0))) \, ds,$$

(A8)

where $K$ is a compact absorbing subset of $D$. Then the condition $\bar{q}_2 < 0$ provides a Bendixson criterion in $D$. As a result, the following theorem holds.

**Theorem A.1** Assume that there exists a compact absorbing set $K \subset D$ and the system (A5) has a unique equilibrium point $X^*$ in $D$. Then $X^*$ is globally asymptotically stable in $D$ if $\bar{q}_2 < 0$.

Appendix 3. A lemma on a special matrix determinant

Let us consider the following square matrix of dimension $n \times n$ ($n \geq 1$):

$$X_1 = \begin{bmatrix}
N_1P_1 & N_1P_2 & \cdots & N_1P_n \\
N_2P_1 & N_2P_2 & \cdots & N_2P_n \\
\vdots & \vdots & \ddots & \vdots \\
NnP_1 & NnP_2 & \cdots & NnP_n
\end{bmatrix},$$

(A9)

where $N_i$ and $P_i$ ($1 \leq i \leq n$) are arbitrary numbers.

**Lemma A.1** The characteristic polynomial of the matrix in Equation (A9) is

$$\det(\lambda I_n - X_1) = \lambda^{n-1} \left( \lambda - \sum_{i=1}^n N_i P_i \right).$$

(A10)

The result follows directly from the trace of $X_1$, by noting that the rank of $X_1$ is 1.

Appendix 4. Table of model parameters

The definition and numerical values of parameters in our ODE and PDE cholera models are provided in Table A1. Note that $p$ (resp. $y$, $w$ and $d$) represents a person (resp. year, week and day).
Table A1. Model parameters used in cholera models.

| Parameter | Definition | Value | References |
|-----------|------------|-------|------------|
| $N$       | Total population size of host | 12,347 | [22]       |
| $d$       | Natural death rate of human   | (43.5 y)$^{-1}$ | [41]       |
| $\beta_h$ | Direct transmission rate      | 0.075 p$^{-1}$ w$^{-1}$ | [22]       |
| $\beta_e$ | Indirect transmission rate    | $1.1 \times 10^{-4}$ w$^{-1}$ | [22]       |
| $\kappa$  | Half saturation rate          | $10^6$ cells $\cdot$ ml$^{-1}$ | [13]       |
| $\gamma$  | Recovery rate                 | (5 d)$^{-1}$ | [13]       |
| $\sigma$  | Rate of host immunity loss    | (3 y)$^{-1}$ | [24]       |
| $K$       | Maximal carrying capacity     | $2 \times 10^6$ cells $\cdot$ ml$^{-1}$ | Assumed   |
| $\delta$  | Bacterial death rate          | (30 d)$^{-1}$ | [13]       |
| $\xi > 0$ | Shedding rate                 | Varied [cells $\cdot$ ml$^{-1}$ w$^{-1}$] | Assumed   |
| $g \geq 0$| Initial bacterial growth rate  | Varied [w$^{-1}$] | Assumed   |