Disease dynamics in a coupled cholera model linking within-host and between-host interactions

Xueying Wang

ABSTRACT

A new modelling framework is proposed to study the within-host and between-host dynamics of cholera, a severe intestinal infection caused by the bacterium *Vibrio cholerae*. The within-host dynamics are characterized by the growth of highly infectious vibrios inside the human body. These vibrios shed from humans contribute to the environmental bacterial growth and the transmission of the disease among humans, providing a link from the within-host dynamics at the individual level to the between-host dynamics at the population and environmental level. A fast-slow analysis is conducted based on the two different time scales in our model. In particular, a bifurcation study is performed, and sufficient and necessary conditions are derived that lead to a backward bifurcation in cholera epidemics. Our result regarding the backward bifurcation highlights the challenges in the prevention and control of cholera.

1. Introduction

Mathematical modelling of infectious diseases has played a significant role in furthering our understanding of disease dynamics and developing effective prevention and control measures against epidemics. For infectious diseases, there are two typical modelling approaches to study the disease infection, and they are usually considered separately. One is to use epidemiological models to investigate disease transmission on a population level [1], which concerns the between-host dynamics of the diseases. The other is to use immunological models to study the disease at the individual level, which is focused on the within-host dynamics. To link these dynamical processes at different levels, coupled (or nested) modelling approach is employed, and indeed this approach can offer new insight into the host–pathogen interactions (see [3,20] and references therein). The within- and between-host dynamics have been studied for several infectious diseases using the nested models (e.g. [7,9,11–16,25]). However, to the best of our knowledge, no existing work has attempted to address the coupled within- and between-host dynamics of cholera (a severe water-borne gastroenteric disease caused by the bacterium *Vibrio cholerae*), despite the fact that a large body of work has been devoted to its population level modelling (e.g. [4–6,17,21,22,26,28,29,32–36]). It is also worth mentioning that some of these...
between-host models have incorporated both human and environmental transmission pathways with nonlinear disease incidence functions (see, e.g. [10,26,27,28,31] and the references therein).

When vibrios from the contaminated environment (typically, water) are ingested by a person and reach the small intestine, complex biological, chemical and genomic interactions occur that lead to human cholera, whose pathogenicity and aetiology remain a very active research area [8]. Generally speaking, the environmental vibrios undergo evolutions through horizontal gene transfer and the formation of a toxin complex. As a result, highly toxic and infectious vibrios are generated with a large inoculum size, which cause severe diarrhea and are subsequently excreted out of the human body through shedding. Clinical studies [23] have found that these freshly shed vibrios are up to 700-fold in terms of infectivity compared to environmental vibrios, and they remain at this highly infectious state out of the human body for several hours during which time they can be transmitted through person-to-person contacts (e.g. by shaking hands, or consuming foods prepared by household people with dirty/contaminated hands). Unfortunately, our knowledge of the precise mechanisms of the bacterial growth and evolution inside the human body and their impact on cholera epidemics at the population level, and the quantification of such mechanisms towards the design of effective disease control and prevention strategies, remains limited at present.

The goal of this work is to fill in our knowledge gap through mathematical modelling and analysis, as a first attempt, in that direction. A multi-scale model is formulated in this paper that couples the within-host and between-host dynamics of cholera. Essentially, we quantitatively distinguish the two types of vibrios: the environmental vibrios which have relatively low infectivity, and the human vibrios (which are formed within human body and then shed out into the environment) that are highly infectious. The highly infectious vibrios live only for a short period of time (in hours) compared to that of the environmental vibrios (in months), but they directly determine the transmission rate through person-to-person contacts and largely contribute to the growth of the vibrios in the aquatic environment. We note that a standard assumption in almost all the afore-mentioned mathematical cholera studies is that the persistence of environmental vibrios and the disease spread among human populations are independent of within-host disease dynamics, which leads to relatively simple, usually constant, formulations in terms of the associated per capita rates of interactions. In contrast, our modelling framework incorporates a general representation of the incidence and the bacterial growth and persistence, and explicitly link disease dynamics at the population and individual levels through the concept of human vibrios. Our approach is different from traditional coupling principles [14] employed in nested models. Specifically, our explicit linkage is achieved by postulating the interaction between the cholera strains in the environment and those inside human hosts, instead of the cell–pathogen immunological interaction within a single host.

We shall mention that in the work of Hartley et al. [17], the authors also classified the vibrios into two types: lower infective and hyper infective. However, there are important differences between these two studies. First, their classification is only concerned with environmental vibrios and there are no within-host dynamics in their work. Second, our model has incorporated both indirect transmission and direct transmission pathways, while the work in [17] only has the indirect transmission route. Third, all the transmission rates in
are constant whereas our model utilizes variable transmission rates, dependent on both the environmental and human vibrios, to explicitly link the fast and slow dynamics. Additionally, their cholera study has a major focus on the medical aspects, whereas our work is more mathematically sophisticated.

The remainder of the paper is organized as follows. Section 2 presents the formulation of the cholera model linking the within-host and between-host dynamics. In Section 3, we analyse the coupled model by using the fast-slow analysis, utilizing the time scale separation for the dynamical processes associated with the individual host and the environment and human population. Global dynamics of the subsystems (i.e. the slow and fast systems) are studied separately. For the slow system, we calculate the corresponding type reproduction number of the infectious humans, and use it to investigate the disease threshold dynamics and analyse the local and global stabilities of the equilibrium solutions. One of our main findings is that our model exhibits a backward bifurcation, which has never been reported in previous cholera studies. A precise condition under which the backward bifurcation occurs in the slow system is obtained. For the fast system, the global stability of the unique equilibrium is established. In addition, we present numerical simulation results to verify our analysis, particularly for the coupled system dynamics. Finally, we wrap up the paper with conclusions and discussion in Section 4.

2. Model formulation

Cholera infection involves different time scales. The between-host and environmental dynamics usually take place on a much slower time scale than that of the within-host dynamics [2]. Here we consider a model system that couples the between-host and within-host dynamics and the interaction between humans and the environment. We use a susceptible-infected-recovered-susceptible epidemic framework [21,34] to describe the between-host dynamics:

\[
\frac{dS}{dt} = \mu N - \beta_H(Z)SI - \beta_L \frac{SB}{\kappa + B} - \mu S + \sigma R,
\]

\[
\frac{dI}{dt} = \beta_H(Z)SI + \beta_L \frac{SB}{\kappa + B} - (\gamma + \mu)I,
\]

\[
\frac{dR}{dt} = \gamma I - (\mu + \sigma)R,
\]

where \(S, I\) and \(R\) refer to the number of susceptible, infected and recovered humans, respectively, \(B\) is the concentration of the bacterium \(Vibrio cholerae\) in the contaminated water, and \(Z\) represents the vibrio concentration in humans. The incidence in this model is represented by both the direct (or person-to-person) transmission, which depends on the human vibrio concentration, and the indirect (or environment-to-person) transmission, which is subject to saturation. The model (1) characterizes the disease dynamics at the population level. The environmental dynamics are represented by the following equation that depicts the evolution of bacteria in the environment:

\[
\frac{dB}{dt} = \xi(Z)I - \delta B,
\]

(2)
where $\xi(Z)$ is the shedding rate that depends on the human vibrios. $\delta$ is the degradation rate of $B$ that includes natural death and other types of removal of pathogens from the aquatic environment. The dynamics at the within-host level are governed by

$$\frac{dZ}{dt} = \frac{1}{\epsilon}(u(B) + v(Z) - \zeta Z), \quad (3)$$

where $0 < \epsilon \ll 1$ is a constant emphasizing the faster time scale compared to that of the between-host dynamics. The environmental vibrios ingested into the human body are transformed into human vibrios at a rate of $u(B)$. Meanwhile, we introduce a function, $v(Z)$, to describe the intrinsic growth of the human vibrios in the course of the biochemical and genomic interactions. In addition, we assume that the human vibrios are removed at a rate of $\zeta Z$ due to shedding or bacterial death.

In our model, the direct human-to-human transmission rate $\beta_H$, the shedding rate of infectious hosts $\xi$, and the human vibrio intrinsic growth rate $v$ are all dependent on $Z$, reflecting the coupling of the within-host and between-host disease dynamics and the link between the fast and slow time scales. The vibrio transition rate $u(B)$ is determined by the environmental vibrio concentration. We make the following biologically feasible assumptions regarding these four functions:

(H1) (a) $\beta_H(0) \geq 0$; (b) $\beta_H'(Z) \geq 0$;
(H2) (a) $\xi(0) \geq 0$; (b) $\xi'(Z) \geq 0$;
(H3) (a) $u(0) = 0$; (b) $u'(B) > 0$;
(H4) (a) $v(0) \geq 0$; (b) $\sup\{v(Z) : Z \geq 0\} \leq v_M$ for some $v_M > 0$; (c) $0 \leq v'(Z) < \zeta$.

These assumptions ensure that all these functions are non-negative and non-decreasing; particularly, the vibrio transition rate $u$ is strictly increasing with respect to $B$. Additional regulations on the function $v(Z)$ in (H4) indicate that, in the absence of the environmental vibrios, the human vibrios will eventually decay to zero inside the human body. Moreover, the definition of all the model parameters is provided in Table A1.

3. Analysis of the disease dynamics

Because of the multiple time scales presented, we will analyse the full system (1)–(3) by using the fast-slow analysis. This approach separates the time scales, and treats the fast system at its quasi-steady state (to which it quickly converges) in the slow system analysis. Readers interested in learning more about fast-slow analysis may refer to [7,11,12] and references therein.

3.1. The slow system

Since $\epsilon \ll 1$, the within-host dynamics (3) can be written as

$$\epsilon \frac{dZ}{dt} = u(B) + v(Z) - \zeta Z. \quad (4)$$

Setting $\epsilon = 0$, we obtain

$$u(B) + v(Z) - \zeta Z = 0. \quad (5)$$
By assumptions (H3)–(H4), differentiating (5) with respect to \( B \) yields
\[
Z'(B) = \frac{u'(B)}{\zeta - v'(Z)} > 0.
\]

So, there exists a unique strictly monotone function
\[
Z = h(B),
\]

such that Equation (5) is satisfied. Consider the associated slow system
\[
\begin{align*}
\frac{dS}{dt} &= \mu N - \beta_H(h(B))SI - \beta_L\frac{SB}{\kappa + B} - \mu S + \sigma R, \\
\frac{dI}{dt} &= \beta_H(h(B))SI + \beta_L\frac{SB}{\kappa + B} - (\gamma + \mu)I, \\
\frac{dR}{dt} &= \gamma I - (\mu + \sigma)R, \\
\frac{dB}{dt} &= \xi(h(B))I - \delta B.
\end{align*}
\]

It is easy to show that the slow system (7) always has the disease-free equilibrium (DFE), \((S, I, R, B) = (N, 0, 0, 0)\).

We assume that the aquatic environment is regarded as a reservoir of the pathogen for the infection. The free-living vibrios adapt to the environment through physiological and genetic changes that can promote their survival and growth [8]. Thus, our model assumes that secondary, new vibrios can be introduced through pathogen shedding by infectious human individuals. Let us denote \( h_0 = h(0) \). By our assumptions, matrices \( F \) and \( V \) associated with the next generation matrix technique [30] take the following forms:
\[
\begin{align*}
F_h &= \begin{bmatrix} N\beta_H(h_0) & N\beta_L/\kappa \\ \xi(h_0) & 0 \end{bmatrix} \quad \text{and} \quad V_h = \begin{bmatrix} \gamma + \mu & 0 \\ 0 & \delta \end{bmatrix}.
\end{align*}
\]

The next generation matrix is
\[
K_h = F_h V_h^{-1} = \begin{bmatrix} N\beta_H(h_0) & N\beta_L \\ \gamma + \mu & \delta \kappa \\ \xi(h_0) & 0 \\ \gamma + \mu \end{bmatrix}.
\]

Thus, by the standard next generation matrix technique [30], the basic reproduction number for the epidemiological dynamics (4) (including between-host and host-environment dynamics) is given by
\[
R_{0b} = \rho(K_h) = \frac{1}{2} \left[ \frac{N\beta_H(h_0)}{\gamma + \mu} + \sqrt{\left( \frac{N\beta_H(h_0)}{\gamma + \mu} \right)^2 + 4 \frac{N\beta_L \xi(h_0)}{\delta \kappa (\gamma + \mu)}} \right],
\]

where \( \rho \) denotes the spectral radius of a square matrix. We comment that in the current paper as well as the work of Wang et al. [34,35], bacterial shedding is interpreted as a
means of generating new pathogen that leads to new infection, whereas in a few other studies [26,28,29,32], pathogen shedding is treated as a transfer between infectious compartments which would result in a different, but equivalent, basic reproduction number. Here we choose to address these two different perspectives through the introduction of the type reproduction number, given below.

We write

$$T_{1h} = \frac{N}{\mu + \gamma} \left( \beta_H(h_0) + \beta_L \frac{\xi(h_0)}{\delta} \right).$$  \hspace{1cm} (11)

$T_{1h}$ is the type reproduction number for the infectious persons, which defines the expected number of secondary infective persons caused by one infectious person in a completely susceptible population [18,24]. More details of the type reproduction number can be found in Appendix 1. Furthermore, it has been shown in [24] that

$$R_{0,b} < 1(= 1, > 1) \iff T_{1h} < 1(= 1, > 1).$$  \hspace{1cm} (12)

### 3.1.1. Equilibrium solutions.

In this section, we will study equilibrium solutions and their stability of the slow system (7). It is easy to verify that Equation (7) always has a DFE, and its endemic equilibrium (EE) satisfies

$$N = S + I + R,$$

$$R = \frac{\gamma}{\mu + \sigma} I,$$

$$I = \frac{B}{\xi(h(B))/\delta},$$

$$S = \frac{(\mu + \gamma)I}{\beta_H(h(B))I + \beta_L B/(B + \kappa)}.$$  \hspace{1cm} (16)

For simplicity, we write

$$p(B) = \frac{\xi(h(B))}{\delta}, \quad q(B) = \beta_H(h(B)).$$  \hspace{1cm} (17)

Substituting Equations (14) and (15) into Equation (13) yields

$$S = \phi(I) := N - \alpha I \quad \text{with} \quad \alpha = 1 + \gamma/\left(\mu + \sigma\right).$$

By Equations (15) and (16), we find that

$$S = f(B) := \frac{\mu + \gamma}{g(B)},$$

where

$$g(B) = q(B) + \frac{\beta_L B}{I(B + \kappa)} = q(B) + \frac{\beta_L p(B)}{B + \kappa},$$  \hspace{1cm} (18)

for which the second equality is obtained by Equation (15). We further assume that

(Ha) There is a constant $\xi_M > 0$ such that $0 \leq \xi(Z) \leq \xi_M$ for all $Z \geq 0$;
(Hb) \( p'(B) \geq 0 \) and \( p''(B) < 0 \);
(Hc) \( q'(B) \geq c_1 \) for some \( c_1 > 0 \), and \( q''(B) \leq -c_2 \) with \( c_2 := 2\beta_L \xi_M / \delta \kappa^3 > 0 \).

These conditions set an upper bound for the maximum rate of shedding that a given human population can contribute to the environmental vibrios, and regulate the shedding function \( \xi(h(B)) \) and the direction transmission function \( \beta_{H}(h(B)) \) as non-decreasing but subject to saturation effects. We do impose slightly stronger conditions on \( \beta_{H}(h(B)) \) than that for \( \xi(h(B)) \), partly due to the fact that the direct transmission rate is typically very sensitive to system dynamics in a cholera model [21].

Let \( I = \tau(B) := B / p(B) \). Since \( p''(B) < 0 \), we have
\[
p(B) - B p'(B) > 0
\] (19)
(which is due to \( 0 \leq p(0) = p(B) - B p'(B) + p''(\theta_B)B^2 / 2 < p(B) - B p'(B) \) for some \( 0 < \theta_B < B \)). It implies that
\[
\tau'(B) = \frac{p(B) - B p'(B)}{(p(B))^2} > 0,
\] (20)
i.e. \( I = \tau(B) \) is strictly increasing function of \( B \). Thus, it has an inverse function (that is strictly monotonic increasing as well), denoted as \( B = w(I) \), such that \( I = \tau(w(I)) \).

The fourth assumption we make is:

(Hd) \( w''(I) \leq 0 \).

Let
\[ \psi(I) = f(w(I)) \]

**Lemma 3.1:** Under assumptions (Ha)–(Hd), \( \psi'(I) < 0 \) and \( \psi''(I) > 0 \).

**Proof:** By direct calculation, we have
\[
g'(B) = q'(B) + \beta_L \frac{p'(B)}{B + \kappa} - \beta_L \frac{p(B)}{(B + \kappa)^2},
\] (21)
\[
g''(B) = q''(B) + \beta_L \frac{2p(B)}{(B + \kappa)^3} + \frac{\beta_L}{B + \kappa} (p''(B) - 2p'(B)/(B + \kappa)).
\] (22)

In view of assumption (Ha)
\[
\lim_{B \to \infty} \frac{p(B)}{(B + \kappa)^2} = 0.
\]
Thus, there exists \( B_M > 0 \) such that \( 0 \leq \beta_L (p(B)/(B + \kappa)^2) \leq c_1 / 2 \) for \( B \geq B_M \). Applying assumptions (Hb)–(Hc) to Equation (21) yields
\[
g'(B) \geq c_1 + \beta_L \frac{p'(B)}{B + \kappa} - \frac{c_1}{2} > 0 \quad \forall B \geq B_M.
\] (23)

Applying assumptions (Ha)–(Hc) to Equation (22), we obtain
\[
g''(B) \leq q''(B) + \beta_L \frac{2\xi_M}{\delta \kappa^3} + \frac{\beta_L}{B + \kappa} (p''(B) - 2p'(B)/(B + \kappa)) \leq 0.
\] (24)
Then, Equations (23) and (24) imply that

$$g'(B) > 0 \quad \text{for all } B.$$  \hfill (25)

On the other hand,

$$\psi'(I) = -\frac{\mu + \gamma}{g(w(I))} (g'(w(I))w'(I)),$$

$$\psi''(I) = -\frac{\mu + \gamma}{g(w(I))} [g''(w(I))(w'(I))^2 + g'(w(I))w''(I)] + \frac{2(\mu + \gamma)}{g(w(I))^3} (g'(w(I))w'(I))^2.$$

It follows from Equations (24), (25), $w'(I) > 0$ and assumption (Hd) that $\psi'(I) < 0$ and $\psi''(I) > 0$. □

The intersections of the curves $S = \phi(I)$ and $S = \psi(I)$ in the interior of $\mathbb{R}^2_+$, determine the nontrivial equilibria. Note that the curve $S = \phi(I)$ is a decreasing line. Its zero is achieved at $I_1 = N/\alpha$. Meanwhile, $\psi(I)$ is strictly decreasing and concave upward by Lemma 3.1. Moreover,

$$\psi(I_1) = \frac{\mu + \gamma}{q(B_1) + \beta_L p(B_1)/(B_1 + \kappa)} > 0 = \phi(I_1),$$  \hfill (26)

with $B_1 = w(I_1)$. At $I = 0$, we see that $\phi(0) = N$ and

$$\psi(0) = \frac{\mu + \gamma}{\beta_H (h_0) + \beta_L \xi(h_0)/(\delta \kappa)} = \frac{N}{T_{1h}}.$$

To study the nontrivial equilibrium solutions of Equation (7), let us consider all three scenarios in terms of $T_{1h}$. Case 1: $T_{1h} > 1$. Then $\psi(0) < \phi(0)$. Lemma 3.1 and Equation (26) imply that these two curves have a unique intersection in the interior $\mathbb{R}^2_+$. Case 2: $T_{1h} = 1$. We have $\psi(0) = \phi(0)$. In this case, these two curves have no (resp. a unique) intersection in the interior $\mathbb{R}^2_+$ when $\psi'(0) \geq -\alpha$ (resp. $\psi'(0) < -\alpha$). Case 3: $T_{1h} < 1$. Then $\psi(0) > \phi(0)$ and there will be three possibilities: (a) these two curves will have no intersection in $\mathbb{R}^2_+$ if $\psi(I) > \phi(I)$ for $I \geq 0$; (b) there will be one intersection if there exists $I^* > 0$ such that $\psi(I^*) = \phi(I^*)$ and $\psi'(I^*) = \phi'(I^*) = -\alpha$; (c) otherwise, there would be two intersections. It shows that the system (7) has at most three biologically feasible equilibria: if $T_{1h} > 1$, this system has two equilibria: DFE and EE; if $T_{1h} < 1$, it can have one, two or three equilibria (and one of them is the DFE) depending on the parameter values. This result indicates that the system (7) may undergo a backward bifurcation.

### 3.1.2. Local stability and bifurcation analysis.

We now want to study the condition that generates a backward bifurcation. At the nontrivial equilibrium solutions of Equation (7), $\phi(I) = \psi(I)$; i.e.

$$N - \alpha I = \psi(I).$$  \hfill (27)
Note that $T_{1h}$ is a scalar multiple of $N$ (see Equation (11)). Let us treat $N$ as the independent variable and fix the rest of model parameters. Differentiating Equation (27) with respect to $N$ yields

$$1 = (\alpha + \psi'(I)) \frac{dI}{dN} =: H(I) \frac{dI}{dN}. \tag{28}$$

Hence, if $H(I) > 0$, $dI/dN > 0$; whereas if $H(I) < 0$, $dI/dN < 0$. One can verify that

$$H'(I) = \psi''(I) > 0. \tag{29}$$

Suppose $N = N_c$ at $T_{1h} = 1$. Together with the existence of a unique EE when $T_{1h} > 1$, we find that, at $I = 0$ with $N = N_c$ (i.e. $T_{1h} = 1$), a backward bifurcation occurs when $H(0) < 0$. In this case (i.e. $H(0) < 0$), there must have a unique $I_s$ such that $H(I) < 0$ for $0 < I < I_s$, and $H(I) > 0$ for $I > I_s$. Thus, $dI/dT_{1h} < 0$ for $0 < I < I_s$, and $dI/dT_{1h} > 0$ for $I > I_s$. Following immediately from the fact that a backward bifurcation occurs at $T_{1h} = 1$ with $I = 0$ when $H(0) < 0$, we obtain the following result:

**Lemma 3.2:** If $\psi'(0) < -\alpha$, then the system (7) undergoes a backward bifurcation at $T_{1h} = 1$ and $I = 0$.

Indeed, if $\psi'(0) > -\alpha$ (i.e. $H(0) > 0$), Equation (29) implies that $H(I) > 0$ and, consequently, $dI/dN > 0$, for all $I \geq 0$. Thus, the system (7) has a forward bifurcation that occurs at $T_{1h} = 1$ and $I = 0$. Therefore, Lemma 3.2 establishes the necessary and sufficient condition under which a backward bifurcation occurs.

Figure 1 illustrates the behaviour of the backward bifurcation for Equation (37). It plots the equilibrium of infected human hosts vs. $T_{1h}$, when $\psi'(0) < -\alpha$. Let $(T_{1s}, I_s)$ be the saddle-node bifurcation point (or turning point), where the two nontrivial equilibrium solutions come together and annihilate each other. As one can see from this figure that the

![Figure 1](image-url)
bifurcation diagram can be divided into three regions vertically: (1) when \(0 \leq T_{1h} < T_{1s}\), the system has a unique equilibrium (i.e. the DFE) that is stable; (2) when \(T_{1s} < T_{1h} < 1\), this system has three equilibria. The top and bottom ones are stable and the middle one is unstable; (3) when \(T_{1h} > 1\), this system has two equilibria and one is stable and the other one is unstable. Here the local asymptotic stability is verified numerically. On the other hand, Figure 1 shows that the backward bifurcation diagram is composed of three branches horizontally: the top, middle and bottom branches. The bottom (resp. top) branch represents the DFE (resp. EE).

We now quantitatively characterize the turning point \((T_{1s}, I_s)\). It is apparent that \(I_s\) can be determined from \(H(I_s) = 0\). To determine \(T_{1s}\), we let \(N_s\) and \(N_c\) denote the value of \(N\) that corresponds to \(T_{1h} = T_{1s}\) and \(T_{1h} = 1\), respectively. Using Equation (27), we have

\[
N_s = \alpha I_s + \psi(I_s) = \alpha I_s + \frac{\mu + \gamma}{\beta_H(h(w(I_s))) + \beta_L \frac{\xi(h(w(I_s))))}{(w(I_s) + \kappa)\delta}}.
\]

Hence,

\[
T_{1s} = \frac{N_s}{\mu + \gamma} \left[\beta_H(h_0) + \beta_L \frac{\xi(h_0)}{\delta\kappa}\right] = \frac{N_s}{N_c}.
\]

By the relationship (12) and Theorem 2 of van den Driessche and Watmough [30], the DFE is locally asymptotically stable for \(T_{1h} < 1\) and unstable for \(T_{1h} > 1\), which resolves the local stability of the bottom branch. The following result is concerned with the local stability of the middle and top branches.

**Theorem 3.3:**

1. The middle branch of the equilibrium solutions of the system (7) in the backward bifurcation diagram is always unstable.
2. The top branch of the equilibrium solutions of the system (7) is locally asymptotically stable provided that \(\gamma \delta \leq (\mu + \sigma)(\mu + \sigma + \gamma)\).

**Proof:** Note that \((S + I + R)(t) = N\) is a constant. Removing the first equation of Equation (7) and replacing \(S\) by \(N - I - R\), we can rewrite the system (7) as an equivalent system:

\[
\begin{align*}
\frac{dI}{dt} &= (N - I - R) \left( q(B)I + \beta_L \frac{B}{\kappa + B} \right) - (\gamma + \mu)I, \\
\frac{dR}{dt} &= \gamma I - (\mu + \sigma)R, \\
\frac{dB}{dt} &= \delta(p(B)I - B),
\end{align*}
\]

where \(p(B)\) and \(q(B)\) are defined in Equation (17).
Let $p = (S, I, R, B)$ denote an equilibrium solution of Equation (7). Linearizing Equation (32) at $p$, we find that the corresponding Jacobian matrix, $J = [J_{ij}]$, takes the form

$$
\begin{bmatrix}
q(B)S - q(B)I - \beta L \frac{B}{B + \kappa} - (\gamma + \mu) & -q(B)I - \beta L \frac{B}{B + \kappa} & q'(B)SI + \beta L S \frac{\kappa}{(B + \kappa)^2} \\
\gamma & -(\mu + \sigma) & 0 \\
\delta p(B) & 0 & \delta(p'(B)I - 1)
\end{bmatrix}.
$$

It follows from direct calculation that the characteristic equation associated with $J$ is

$$
\lambda^3 + a\lambda^2 + b\lambda + c = 0, \tag{34}
$$

for which

$$
a = -(J_{11} + J_{22} + J_{33}),
$$

$$
b = J_{11}J_{22} + J_{12}J_{33} - J_{12}J_{21} + (J_{11}J_{33} - J_{13}J_{31}),
$$

$$
c = J_{12}J_{21}J_{33} - J_{22}(J_{11}J_{33} - J_{13}J_{31}).
$$

We now prove part 1. Assume that this system has a backward bifurcation. To show the middle branch in the backward bifurcation diagram is unstable, let us consider the equation that solves the nontrivial equilibrium solutions in terms of $B$; that is, $\phi(I(B)) = \psi(I(B))$ for which $I(B) = B/P(B)$. This equation gives

$$
N - \alpha I(B) = \frac{\mu + \gamma}{g(B)}. \tag{36}
$$

Multiplying $g(B)$ on both sides of Equation (36) yields

$$
g(B)(N - \alpha I(B)) = \mu + \gamma. \tag{37}
$$

Regard $N$ as an independent variable and fix the other model parameters. Differentiating Equation (37) with respect to $N$, we obtain

$$
g'(B)(N - \alpha I(B)) \frac{dB}{dN} + g(B) \left[ 1 - \alpha I'(B) \frac{dB}{dN} \right] = 0. \tag{38}
$$

This yields

$$
G(B) \frac{dB}{dN} = g(B) > 0, \tag{39}
$$

where

$$
G(B) = \alpha g(B)I'(B) - g'(B)[N - \alpha I(B)]. \tag{40}
$$

Applying Equations (20) and (21), we have

$$
BG(B) = \alpha \frac{p(B) - Bp'(B)}{p(B)} \left[ q(B) \frac{B}{p(B)} + \beta L \frac{B}{B + \kappa} \right] - SBq'(B) - S\beta L \frac{Bp'(B)}{B + \kappa} + S\beta L \frac{Bp(B)}{(B + \kappa)^2}. \tag{41}
$$
with $S = N - \alpha I(B)$. Hence, Equation (39) implies that $BG(B)$ and $dB/dN$ have the same sign. By definition $\alpha = 1 + \gamma/\mu + \sigma$, it yields

$$\gamma + \mu + \sigma = \alpha(\mu + \sigma).$$

(42)

Direct calculation together with Equation (42) leads to

$$c = J_{12}J_{21}J_{33} - J_{22}(J_{11}J_{33} - J_{13}J_{31}) = \delta(\mu + \sigma)BG(B).$$

(43)

This implies that $c$ has the same sign as $dB/dN$. Particularly, in the case where the system (7) has a backward bifurcation, there exists a unique $B_s > 0$, at which $I(B_s) = I_s$, such that $dB/dN < 0$ if $0 < B < B_s$; whereas $dB/dN > 0$ if $B > B_s$. The middle branch of equilibrium solution (on which $dB/dN < 0$) of a backward bifurcation satisfies $c < 0$, and hence it must be unstable.

We now verify part 2. Each point on the top branch of the equilibrium solutions is an EE. In what follows, we denote the EE by $(S, I, R, B)$ and all the arguments below are restricted to this point. According to the Routh–Hurwitz criterion, it suffices to prove that

$$a > 0, \quad b > 0, \quad c > 0, \quad ab - c > 0.$$  

(44)

Clearly, $c > 0$ since $dB/dN > 0$ along the EE. On the other hand, (a) $q(B)S - (\gamma + \mu) = -\beta_LSB/(I(B + \kappa))$; (b) $I = B/p(B)$. Hence, we can rewrite $J_{11}$ and $J_{33}$ as follows:

$$J_{11} = -\left( q(B)I + \beta_L \frac{B}{B + \kappa} + \beta_L S \frac{B}{I(B + \kappa)} \right),$$

$$J_{33} = \frac{\delta}{p(B)}(p'(B)B - p(B)).$$

(45)

It is obvious that $J_{11} < 0$ and $J_{33} < 0$ (where the second inequality follows immediately from Equation (19)). Meanwhile, by (Hb)–(Hc), one can verify that the following holds: $J_{12} < 0, J_{13} > 0, J_{21} > 0, J_{22} < 0$, and $J_{21} < 0$. Thus, $a > 0$.

To prove Equation (44), it remains to show that $b > 0$ and $ab - c > 0$. Again by Equation (42), we obtain

$$J_{11}J_{22} + J_{22}J_{33} - J_{12}J_{21} = (\mu + \sigma)\alpha \left[ q(B)\frac{B}{p(B)} + \beta_L \frac{B}{B + \kappa} \right]$$

$$+ (\mu + \sigma)\beta_L \frac{Sp(B)}{B + \kappa} + \delta(\mu + \sigma)\frac{p(B) - Bp'(B)}{p(B)}.$$  

(46)

Additionally,

$$\frac{1}{\delta}(J_{11}J_{33} - J_{13}J_{31}) = \frac{p(B) - Bp'(B)}{p(B)} \left[ q(B)\frac{B}{p(B)} + \beta_L \frac{B}{B + \kappa} \right]$$

$$- SBq'(B) - S\beta_L \frac{Bp'(B)}{B + \kappa} + S\beta_L \frac{Bp(B)}{(B + \kappa)^2}.  \quad (47)$$
In view of Equations (39) and (47), we have
\[
\delta[BG(B) - (J_{11}J_{33} - J_{13}J_{31})/\delta] = \delta \left( 1 - \frac{Bp'(B)}{P(B)} \right) (\alpha - 1) \left[ q(B) \frac{B}{p(B)} + \beta_L \frac{B}{B + \kappa} \right] \\
\leq \delta \frac{\gamma}{\mu + \sigma} \left[ q(B) \frac{B}{p(B)} + \beta_L \frac{B}{B + \kappa} \right] \\
\leq J_{11}J_{22} - J_{12}J_{21}. \quad (48)
\]
This first inequality of Equation (48) holds because of Equation (20). The second inequality of Equation (48) follows from Equation (46) and \( \gamma \delta \leq (\mu + \sigma)(\mu + \sigma + \gamma) \). Thus, Equation (48) shows that \( b > 0 \). On the other hand, we have
\[
ab - c > (-J_{22}) \delta BG(B) - c = (\mu + \sigma) \delta BG(B) - c = 0. \quad (49)
\]
This completes the proof. \( \blacksquare \)

### 3.1.3. Global stability analysis.

Let \( \Omega = \{(S, I, R, B) \in \mathbb{R}^4 : S, I, R \leq N, B \leq B_M := \xi_M N/\delta \} \). It is easy to see that \( \Omega \) is forward invariant for model (7), which can be verified by showing the solutions of Equation (7) do not cross the boundary of \( \Omega \). Write \( \beta_H^M = \max_{0 \leq B \leq B_M} \{\beta_H(h(B))\} \). The following results (i.e. Theorems 3.4 and 3.5) establish global dynamics of the slow system (7).

**Theorem 3.4:** If \( T_{1h}^M := (N/(\mu + \gamma))(\beta_H^M + \beta_L(\xi_M/\delta \kappa)) < 1 \), then the DFE of the slow system (7) is globally asymptotically stable in the region \( \Omega \).

**Proof:** The idea of the proof is to construct a Lyapunov function. Let \( w = (\beta_H^M N, \beta_L N/\kappa) \),
\[
F_M = \begin{bmatrix} \beta_H^M N & \beta_L N/\kappa \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad V_M = \begin{bmatrix} \mu + \gamma & 0 \\ -\xi_M & \delta \end{bmatrix}.
\]
One can verify directly that \( T_{1h}^M = \rho(F_M V_M^{-1}) = \rho(V_M^{-1} F_M) \) and \( w V_M^{-1} F_M = T_{1h}^M w \). Define a Lyapunov function as \( Q = w V_M^{-1} Y \) with \( Y = (I, B)^T \). Differentiating \( Q \) along the solutions of Equation (7) and applying the definition of \( \beta_H^M \) and \( \xi_M \), we obtain
\[
Q' = w V_M^{-1} \frac{dY}{dt} \leq (F_M - V_M) Y = (T_{1h}^M - 1) w Y. \quad (50)
\]
If \( T_{1h}^M < 1 \), then \( Q' = 0 \) implies that \( w Y = 0 \), and hence \( I = B = 0 \). Furthermore, by the first and third equation of the slow system (7), \( S = N \) and \( R = 0 \). Thus, the only invariant set in \( \Omega \) where \( Q' = 0 \) is the singleton \( \{(N, 0, 0, 0)\} \). It shows that the DFE is globally asymptotically stable in \( \Omega \) when \( T_{1h}^M < 1 \). \( \blacksquare \)

**Remark of Theorem 3.4:** Consider the case of the backward bifurcation. By Equation (31) and the definition of \( T_{1h}^M \), we find that
\[
\frac{1}{T_{1s}} = \frac{N_S}{N_C} < \frac{\beta_H(h(w(I_s))) + \beta_L w(I_s) + \kappa)}{\beta_H(h_0) + \beta_L w(h_0) + \kappa)} < \frac{\beta_H + \beta_L \xi_M}{\beta_H(h_0) + \beta_L \xi(h_0)/\delta \kappa} = \frac{T_{1h}^M}{T_{1h}}. \quad (51)
\]
Thus, when \( T_{1h}^M < 1 \), we have \( T_{1h}/T_{1s} < 1 \); i.e. \( T_{1h} < T_{1s} \), and the DFE is globally asymptotically stable in the backward bifurcation domain. \( \blacksquare \)
To study the global stability of the EE, we introduce another assumption:

(C1) \( \sup_{\Omega} \{Sq(B)\} \leq \min\{\gamma/2 - \sigma, \sigma + \mu/2\} \);  
(C2) \( \sup_{\Omega} \{\delta p'(B)I\} \leq \mu/2 + \sigma \).

These two conditions set some further regulation on the human shedding and direct transmission functions to ensure the global stability.

**Theorem 3.5:** If \( T_1h > 1 \), the EE of Equation (7) is globally asymptotically stable in \( \Omega^0 \) provided that (C1)–(C2) hold additionally.

**Proof:** If \( T_1h > 1 \), Theorem 3.3 implies that the slow system (7) has a unique EE. In what follows, we will employ the geometric approach (based on the second additive compound matrix) [19] to verify the global stability of the EE (see Appendix 3 for an outline of the geometric approach). Since the DFE is unstable and it lies on the boundary of the domain \( \Omega \), the disease is uniformly persistent in \( \Omega^0 \); i.e.

\[
\liminf_{t \to \infty} \| (I(t), R(t), B(t)) \| > a_0,
\]

for some \( a_0 > 0 \), where \( \| \cdot \| \) is the \( l_2 \) vector norm in \( \mathbb{R}^3 \). Moreover, \( \Omega \) is compact. This implies that there must exist a compact absorbing set in \( \Omega \). According to Theorem A.1 of Li and Muldowney [19], it remains to verify that the generalized Bendixson criterion \( \tilde{q}_2 < 0 \) (see Appendix 3 for the definition of \( \tilde{q}_2 \)). The idea of the proof is to construct a norm in \( \mathbb{R}^3 \) and a matrix-valued function \( P(S, I, B) \) such that \( \tilde{q}_2 < 0 \).

First, by \( R = N - S - I \), we rewritten Equation (7) as

\[
\begin{align*}
\frac{dS}{dt} &= \mu N - S \left( q(B)I + \beta_L \frac{B}{B + \kappa} \right) - \mu S + \sigma (N - S - I), \\
\frac{dI}{dt} &= S \left( q(B)I + \beta_L \frac{B}{B + \kappa} \right) - (\mu + \gamma) I, \\
\frac{dB}{dt} &= \delta p(B) I - \delta B.
\end{align*}
\]

For simplicity, let

\[
\begin{align*}
\theta_1 &= q(B)I + \beta_L \frac{B}{B + \kappa}, \quad \theta_2 = q(B)S, \\
\theta_3 &= \left( q(B)I + \beta_L \frac{\kappa}{(B + \kappa)^2} \right) S, \quad \theta_4 = \delta (p'(B)I - 1).
\end{align*}
\]

By direct calculation, we find that the Jacobian matrix of the linearized system of Equation (52) is

\[
\tilde{J} = \begin{bmatrix}
-\theta_1 - (\mu + \sigma) & -\theta_2 - \sigma & -\theta_3 \\
\theta_1 & \theta_2 - (\mu + \gamma) & \theta_3 \\
0 & \delta p(B) & \theta_4
\end{bmatrix}
\]
and the associated second additive compound matrix is
\[
\tilde{j}^{[2]} = \begin{bmatrix}
-\theta_1 + \theta_2 - (2\mu + \sigma + \gamma) & \theta_3 & \theta_3 \\
\delta p(B) & -\theta_1 - (\mu + \sigma) + \theta_4 & -\theta_2 - \sigma \\
0 & \theta_1 & \theta_2 - (\mu + \gamma) + \theta_4
\end{bmatrix}.
\]
Define
\[
P = \text{diag}\left[1, \frac{I}{B}, \frac{I}{B}\right].
\]
Clearly, \(P\) is nonsingular and \(C^1\) in \(\Omega^0\). Let \(f\) denote the vector field of Equation (7). Then
\[
PfP^{-1} = \text{diag}\left[0, \frac{I'}{I} - \frac{B'}{B}, \frac{I'}{I} - \frac{B'}{B}\right]
\]
and
\[
p\tilde{j}^{[2]}p^{-1} = \begin{bmatrix}
-\theta_1 + \theta_2 - (2\mu + \sigma + \gamma) & \theta_3 & \theta_3 \\
\delta p(B) & -\theta_1 - (\mu + \sigma) + \theta_4 & -\theta_2 - \sigma \\
0 & \theta_1 & \theta_2 - (\mu + \gamma) + \theta_4
\end{bmatrix}.
\]
Rewrite the matrix \(U := PfP^{-1} + p\tilde{j}^{[2]}p^{-1}\) in the block form:
\[
U = \begin{bmatrix}
U_{11} & U_{12} \\
U_{21} & U_{22}
\end{bmatrix},
\]
where
\[
U_{11} = -\left(q(B)I + \beta_L \frac{B}{B + \kappa}\right) + q(B)S - (2\mu + \sigma + \gamma),
\]
\[
U_{12} = \frac{SB}{I} \left(q'(B)I + \beta_L \frac{\kappa}{(B + \kappa)^2}\right) [1 \ 1],
\]
\[
U_{21} = \begin{bmatrix}
\delta p(B) \frac{I'}{B} \\
0
\end{bmatrix},
\]
\[
U_{22} = \begin{bmatrix}
u_{11} & u_{12} \\
u_{21} & u_{22}
\end{bmatrix},
\]
with
\[
u_{11} = -\left(q(B)I + \beta_L \frac{B}{B + \kappa}\right) + \delta(p'(B)I - 1) - (\mu + \sigma) + \frac{I'}{I} - \frac{B'}{B},
\]
\[
u_{12} = -q(B)S - \sigma,
\]
\[
u_{21} = q(B)I + \beta_L \frac{B}{B + \kappa},
\]
\[
u_{22} = q(B)S + \delta(p'(B)I - 1) - (\mu + \gamma) + \frac{I'}{I} - \frac{B'}{B}.
\]
We now choose the following vector norm \(| \cdot |\) in \(\mathbb{R}^3\)

\[|(x_1, x_2, x_3)| = \max(|x_1|, |x_2| + |x_3|).\]

One can verify that the Lozinski\' measure \(\mathcal{M}(U)\) with respect to this norm satisfies

\[\mathcal{M}(U) \leq \sup\{g_1, g_2\},\]

where

\[g_1 = \mathcal{M}_1(U_{11}) + |U_{12}|,\]
\[g_2 = |U_{21}| + \mathcal{M}_1(U_{22}).\]

Here \(|U_{12}|\) and \(|U_{21}|\) denote matrix norms induced by the \(l_1\) vector norm, \(\mathcal{M}_1\) denotes the Lozinski\' measure in terms of the \(l_1\) norm. Specifically,

\[g_1 = U_{11} + |U_{12}|\]
\[= -\left( q(B)I + \beta_L \frac{B}{B + \kappa} \right) + q(B)S - (2\mu + \sigma + \gamma) + \left[ q'(B)I + \frac{\beta_1 \kappa}{(B + \kappa)^2} \right] SB \frac{I}{I},\]

\[g_2 = |U_{21}| + \max\{u_{11} + |u_{21}|, |u_{12}| + u_{22}\}\]
\[= \delta p(B) \frac{I}{B} + \delta(p'(B)I - 1) - (\mu + \sigma) + \frac{I'}{I} - \frac{B'}{B} + \sup\{0, 2(q(B)S + \sigma) - \gamma\}.\]

In view of \(I' = S(q(B)I + \beta_L(B/(B + \kappa)) - (\mu + \gamma)I\), we get

\[-(\mu + \gamma) = \frac{I'}{I} - Sq(B) - \frac{S \beta_L B}{I B + \kappa},\]

and this yields

\[g_1 = -q(B)I - \beta_L \frac{B}{B + \kappa} - (\mu + \sigma) + SBq'(B) + \frac{S\beta_L \kappa}{(B + \kappa)^2} \frac{B}{I} + \frac{I'}{I} - \frac{S \beta_L B}{I B + \kappa}\]
\[= -q(B)I - \beta_L \frac{B}{B + \kappa} - (\sigma + \mu/2) + SBq'(B) - \frac{S\beta_L B}{(B + \kappa)I} \left(1 - \frac{\kappa}{B + \kappa}\right) + \frac{I'}{I} - \frac{\mu}{2}\]
\[\leq \frac{I'}{I} - \frac{\mu}{2}. \quad (53)\]

The last inequality follows from (C1) and \(q(B) \geq Bq'(B)\) for all \(B \geq 0\). Using a similar argument together with (C2), we have

\[g_2 = \delta p'(B)I + \frac{I'}{I} - (\mu + \sigma) + \sup\{0, 2(q(B)S + \sigma) - \gamma\}\]
\[\leq \frac{I'}{I} - \frac{\mu}{2}. \quad (54)\]

Thus, Equations (53) and (54) yield

\[\mathcal{M}(U) \leq \frac{I'}{I} - \frac{\mu}{2}.\]
By \(0 \leq I(t) \leq N\), we obtain that, if \(t\) is large enough,
\[
\frac{\ln(I(t)) - \ln(I(0))}{t} \leq \frac{\mu}{4}.
\]

Therefore,
\[
\frac{1}{t} \int_0^t \mathcal{M}(U) \, dz \leq \frac{1}{t} \int_0^t \left( \frac{I'(z)}{I(z)} - \frac{\mu}{2} \right) \, dz = \frac{\ln(I(t)) - \ln(I(0))}{t} - \frac{\mu}{2} \leq -\frac{\mu}{4},
\]
if \(t\) is large enough. This in turn implies that \(\bar{q}_2 \leq -\mu/4 < 0\) and it completes the proof. \(\blacksquare\)

**Remark on Theorem 3.5:** From the proof of the theorem, we can observe that the conditions (C1)–(C2) can be relaxed to:

**(C1')** \(\sup \Omega \{ SBq'(B) - q(B)I - \beta_L(B/(B + \kappa)) - (S\beta_L(B/(B + \kappa))I)(1 - \kappa/(B + \kappa)) \} \leq \sigma + \mu/2; \) and one of the following:

**(C2')** \(\sup \Omega \{ Sq(B) \} \leq \gamma/2 - \sigma, \sup \Omega \{ \delta p'(B)I \} \leq u/2 + \sigma; \) or

**(C2'')** \(\sup \Omega \{ Sq(B) \} \geq \gamma/2 - \sigma, \sup \Omega \{ 2Sq(B) + \delta p'(B)I \} \leq \gamma + u/2 - \sigma. \) \(\blacksquare\)

### 3.2. The fast system

Introduce a fast timescale \(s = t/\epsilon\). The full system (1), (2) and (3) in terms of the fast time \(s\) can be written as

\[
\frac{dS}{ds} = \epsilon \left( \mu N - \beta_H(Z)SI - \beta_L \frac{SB}{\kappa + B} - \mu S + \sigma R \right),
\]
\[
\frac{dI}{ds} = \epsilon \left( \beta_H(Z)SI + \beta_L \frac{SB}{\kappa + B} - (\gamma + \mu)I \right),
\]
\[
\frac{dR}{ds} = \epsilon (\gamma I - (\mu + \sigma)R),
\]
\[
\frac{dB}{ds} = \epsilon (\xi(Z) - \delta B),
\]
\[
\frac{dZ}{ds} = u(B) + v(Z) - \zeta Z.
\]

Populations of human hosts and bacteria in the environment change very little within the fast time scale; that is, if \(\epsilon \ll 1, S, I, R\) and \(B\) can be regarded as a constant. Setting \(\epsilon = 0\) in Equation (55) leads to the fast system

\[
\frac{dZ}{ds} = u(B) + v(Z) - \zeta Z.
\]

Integrating Equation (56) in terms of \(s\) yields

\[
Z(s) = Z(0) \exp(-\zeta s) + \frac{u(B)}{\zeta} (1 - \exp(-\zeta s))
\]
\[
+ \int_0^s [v(Z(w)) \exp(-\zeta (s - w))] \, dw, \quad s \geq 0.
\]
The following result establishes the existence, uniqueness and global stability of the equilibrium solution of the fast system.

**Theorem 3.6:** For each $B \geq 0$, the fast system (56) has a unique equilibrium and this equilibrium is globally asymptotically stable.

**Proof:** We know that $Z = h(B)$ is the unique solution of $u(B) + v(Z) - \xi Z = 0$. Define $Z_\infty = \lim_{s \to \infty} Z(s)$. We proceed to show that $Z_\infty = h(B)$. One can verify by the standard analysis technique that

$$
\lim_{s \to \infty} \int_0^s [v(Z(w)) \exp(-\xi(s-w))] \, dw = \frac{v(Z_\infty)}{\xi}.
$$

If $s \to \infty$, Equation (57) yields

$$
Z_\infty = \frac{u(B)}{\xi} + \frac{v(Z_\infty)}{\xi}.
$$

This implies that $Z_\infty = h(B)$ is the equilibrium solution and it is globally asymptotically stable. ■

When $B = 0$, the fast system (56) depicts the within-host dynamics in isolation. Particularly, if $v(0) = 0$, then $Z_\infty = 0$; i.e. the system has an infection-free equilibrium (IFE) and the within-host infection dies out. This can be easily obtained by looking at the within-host thresholds. Specifically, let $R_w(B)$ denote the within-host reproduction number, which is a function the bacterial level $B$ in the environment. It is straightforward to derive that $R_w(B) = v'(h(B))/\xi$ and that $R_w(B)$ is an increasing function of $B$ (since $h(B)$ is increasing). Let $R_{0w}$ denote the within-host baseline reproduction number. In view of (H4)

$$
R_{0w} = R_w(0) = \frac{v'(h_0)}{\xi},
$$

and $R_{0w} < 1$. It then follows immediately from Theorem 2 in [30] that the IFE is locally asymptotically stable.

When $B > 0$, we have $Z_\infty > 0$, and the fast system (56) captures the within-host dynamics linked to the between-host dynamics through $u(B)$. In this case, the system has a unique nontrivial equilibrium and the within-host infection persists. Furthermore, the resulting bacterial level in humans will provide feedbacks for the between-host dynamics (involving environment-to-human and human-to-human transmission pathways). It can be observed, however, that this nontrivial equilibrium can be reduced to the IFE of another system through a simple translation, and its stability can then be easily understood. Let $\tilde{Z} = Z - h(B)$. Then

$$
\frac{d\tilde{Z}}{ds} = u(B) + v(\tilde{Z} + h(B)) - \xi(\tilde{Z} + h(B))
= v(\tilde{Z} + h(B)) - v(h(B) - \xi \tilde{Z})
= v'(\theta_{h(B),\tilde{Z}}) \tilde{Z} - \xi \tilde{Z} \quad \text{where } \theta_{h(B),\tilde{Z}} \text{ is between } h(B) \text{ and } \tilde{Z} + h(B)
\leq (\max v' - \xi) \tilde{Z}.
$$

(59)
The second equality of Equation (59) is obtained by using \( u(B) + v(h(B)) - \zeta h(B) = 0 \). Thus, the equilibrium \( Z_\infty \) of the fast system (56) corresponds to the unique IFE \( \tilde{Z} = 0 \) of the system (59) which is, based on assumption (H4), both locally and globally asymptotically stable.

### 3.3. The coupled system (1)–(3)

In this section, we will examine the dynamics of the full system (1)–(3). Let

\[
F = \begin{bmatrix}
N \beta_H(h_0) & N \beta_L/\kappa & 0 \\
\xi(h_0) & 0 & 0 \\
0 & u'(0)/\epsilon & v'(0)/\epsilon
\end{bmatrix}
\quad \text{and} \quad
V = \begin{bmatrix}
\mu + \gamma & 0 & 0 \\
0 & \delta & 0 \\
0 & 0 & \zeta/\epsilon
\end{bmatrix}.
\]

(60)

Similar to our discussion made to the slow system, the matrix \( F \) represents new infections and \( V \) represents transition terms of the coupled system (1)–(3). Following the next generation method [30], the associated basic reproduction number \( R_0 \) is defined as the spectral radius of the next generation matrix \( K = FV^{-1} = (k_{ij}) \). One can verify that

\[
R_0 := \rho(K) = \max\{R_{0b}, R_{0w}\},
\]

(61)

where \( R_{0b} \) and \( R_{0w} \) are the between-host (or epidemiological) and within-host (or individual) thresholds defined in Equations (10) and (58), respectively. Since \( R_{0w} < 1 \),

\[
R_0 > 1 \ (= 1, < 1) \iff R_{0b} > 1 \ (= 1, < 1).
\]

These results indicate that (1) when \( R_{0b} \geq 1 \) or \( R_{0w} \leq R_{0b} < 1, R_0 = R_{0b} \), and the epidemiological reproduction number (i.e. the disease threshold of the slow system (7)) will be the disease threshold of the full system. Particularly, when a forward bifurcation occurs, the disease will persist (resp. die out) if the epidemiological reproduction number is greater (resp. less) than one; (2) when \( R_{0b} < R_{0w} < 1, R_0 = R_{0w} \) and within-host reproduction number will govern the disease threshold dynamics of the full system. On the other hand, it follows from direct calculation that the type reproduction number associated with infectious humans is

\[
T_1 = k_{11} + k_{12}k_{21} = \frac{N}{\mu + \gamma} \left( \beta_H(h_0) + \beta_L \frac{\xi(h_0)}{\delta \kappa} \right) = T_{1h}.
\]

(62)

The relationship in Equation (62) implies that this type reproduction number for the coupled system remains the same as the one for the slow system. This is due to our relatively simple formulation of the within-host dynamics, which leads to a one-to-one correspondence between the bacterial concentration in humans and that in the environment at the equilibrium state. Inclusion of human bacterial process generates a backward bifurcation not only in the coupled system but also in the epidemiological model (i.e. the slow system). The occurrence of the backward bifurcation indicates that an EE can exist and the infection can persist even if the epidemiological reproduction number is reduced to below the unity. This result underscores the challenges in the prevention and control of cholera. A numerical illustration of the backward bifurcation for the coupled system is displayed in Figure 2, which plots the number of infectious hosts as a function of time. It shows that
Figure 2. Illustration of the backward bifurcation in the coupled system (1)–(3). Each curve plots the number of infectious human hosts vs. time (in weeks). The black dotted (resp. blue dash-dot and red solid) curve shows the case where $T_1 < 1$ and the disease dies out (resp. the disease persists when $T_1 < 1$ and $T_1 > 1$). Note that the red curve is continuous and its peak value for the number of infectious hosts is much higher than the window displayed. Hence two disconnected pieces of red curves are shown. In this example, we set $\xi(Z) = \xi_M$, $\beta_H(Z) = b_0 + b_1Z + b_2Z^2$, $u(B) = B$ and $v(Z) = 0$.

(a) the disease can either die out (see the black dotted curve in Figure 2) or persist (see the blue dash-dot curve in Figure 2) even if the type reproduction number for infectious humans $T_1$ is less than one; (b) the full system can undergo a catastrophic transition of the disease in a sense that a small perturbation of $T_1$ from below one to above one can lead to a dramatic increase of disease prevalence. For instance, it is shown in this figure that the prevalence level in terms of the infectious human population size at the equilibrium

Figure 3. Numerical verification of the global stability of the EE associated with the coupled system (1)–(3) when $T_1 > 1$. This figure plots the phase portrait on $S$–$I$ plane. It shows that all the trajectories with biologically meaningful initial conditions converge to the EE (displayed in the black dot).
can increase from 0 to 56 or so, as a matter of the fact that \( T_1 \) is slightly elevated across 1. It is worth mentioning that the initial population size of infectious persons in the case of zero-prevalence is 100, whereas that in the case of 56-prevalence is only 10.

We also numerically verify the global stability of the EE for the full system when \( T_1 > 1 \) by simulating the trajectories with many different initial conditions in the biologically relevant domain. Figure 3 illustrates the projection of these trajectories onto \( S–I \) plane. It shows that the solutions converge to the EE (displayed in black dot) over time, an evidence that the EE is globally asymptotically stable.

### 4. Conclusion and discussion

The main contribution of this work is that we have developed a new mathematical modelling framework to study the within- and between-host dynamics of cholera, and this framework can be easily extended to other infectious diseases with free-living pathogens in the environment. The explicit linkage between disease dynamics at different scales is formulated by presuming a general representation for both the direct transmission and the pathogen shedding, and the interaction between environmental vibrios and human vibrios. Using the proposed model, we have carefully analysed the within- and between-host dynamics of the disease. Our results show that dynamics of the coupled system can be recovered by those of the slow and fast systems obtained by the fast-slow approximations. Particularly, the fast system exhibits a unique equilibrium, either infection-free or endemic, and it is globally asymptotically stable. In contrast, the slow system that captures between-host dynamic is much more complex. This subsystem can undergo a backward bifurcation, and the precise condition for the occurrence of a backward bifurcation is derived. It indicates that reducing the type reproduction number of infectious persons (or the basic reproduction number) below one is not sufficient for eradicating the disease. This finding is distinct from all previous mathematical cholera studies, in which the models only exhibit a forward bifurcation and reducing a disease threshold (e.g. the basic reproduction number) below one leads to disease extinction. The occurrence of the backward bifurcation in our model is a direct result of coupling the within- and between-host cholera dynamics.

There are several limitations in our current study. Our model does not explicitly make a distinction among human hosts, other than partially accounting for it through the ingestion rate \( u(B) \) and intrinsic growth rate \( v(Z) \), and thus each infected individual undergoes the same type of within-host dynamics. This assumption allows our model to be mathematically tractable, though we admit that it is unrealistic and, in reality, many factors (such as age, life style, health condition, and susceptibility level) will play together and possibly lead to very different disease dynamics for different individuals. Another limitation in our current model is the relatively simple formulation of the within-host dynamics which may not be able to capture the full complexity of the biological and pathophysiological processes inside the human body. Incorporation of the immunological modelling principles with detailed pathogen–cell interactions into our model will enrich the within-host dynamics, and lead to potentially deeper understanding of the disease mechanism. The challenge, however, is that the within-host system becomes high dimensional which will complicate the analysis for both the fast and slow system. In addition, our current model assumes that the growth of human vibrios only depends on the environmental vibrios and the intrinsic dynamics of the human vibrios. Practically, the within-host cholera dynamics
may also connect to several epidemiological factors, such as the disease prevalence. We plan to investigate these issues in our future study, using both analytical and numerical approaches.

**Acknowledgments**

The authors would like to thank the editor and the anonymous reviewers for their suggestions that have improved this paper.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was partially supported by a grant from the Simons Foundation (#317047 to X. Wang). J. Wang was partially supported by the National Science Foundation [grant number 1216936], [grant number 1412826] and [grant number 1520672].

**References**

[1] R.M. Anderson and R.M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.

[2] R. Antia, M.A. Nowak, and R.M. Anderson, *Antigenic variation and the within-host dynamics of parasites*, Proc. Natl. Acad. Sci. USA 93 (1996), pp. 985–989.

[3] B. Boldin and O. Diekmann, *Superinfections can induce evolutionarily stable coexistence of pathogens*, J. Math. Biol. 56 (2008), pp. 635–672.

[4] V. Capasso and S.L. Paveri-Fontana, *A mathematical model for the 1973 cholera epidemic in the European Mediterranean region*, Rev. Epidemiol. Sante 27(2) (1979), pp. 121–132.

[5] V. Capasso and G. Serio, *A generalization of the Kermack–McKendrick deterministic epidemic model*, Math. Biosci. 42(1) (1978), pp. 43–61.

[6] A. Carpinter, *Behavior in the time of cholera: Evidence from the 2008–2009 cholera outbreak in Zimbabwe*, in *Social Computing, Behavioral-Cultural Modeling and Prediction*, J. Salerno, S.J. Yang, D. Nau, and S.-K. Chai, eds., Springer, College Park, MD, 2014, pp. 237–244.

[7] X. Cen, Z. Feng, and Y. Zhao, *Emerging disease dynamics in a model coupling within-host and between-host systems*, J. Theor. Biol. 361 (2014), pp. 141–151.

[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. 3–26.

[9] D. Coombs, M.A. Gilchrist, and C.L. Ball, *Evaluating the importance of within- and between-host selection pressures on the evolution of chronic pathogens*, Theor. Popul. Biol. 72 (2007), pp. 576–591.

[10] M.C. Eisenberg, Z. Shuai, J.H. Tien, and P. van den Driessche, *A cholera model in a patchy environment with water and human movement*, Math. Biosci. 246(1) (2013), pp. 105–112.

[11] Z. Feng, J. Velasco-Hernandez, B. Tapia-Santos, and M.C.A. Leite, *A model for coupling within-host and between-host dynamics in an infectious disease*, Nonlinear Dyn. 68 (2012), pp. 401–411.

[12] Z. Feng, J. Velasco-Hernandez, and B. Tapia-Santos, *A mathematical model for coupling within-host and between-host dynamics in an environmentally-driven infectious disease*, Math. Biosci. 241 (2013), pp. 49–55.

[13] Z. Feng, X. Cen, Y. Zhao, and J.X. Velasco-Hernandez, *Coupled within-host and between-host dynamics and evolution of virulence*, Math. Biosci. (2015). Available at http://dx.doi.org/10.1016/j.mbs.2015.02.01.
Appendix 1. 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 [18,24]. It is an extension of the basic reproduction number $R_0$.

The type reproduction number $T_i$ for the control of the target population type $i$ is defined in Refs. [18,24] 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. Here $K$ is the next generation matrix of size $n$. $I$ is the identity matrix. $e_i$ is a column vector with all zero entries except that the $i$th entry is 1.

Write $K = (k_{ij})$. Particularly, if $n = 2$, 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)

$T_1$ (resp. $T_2$) exists provided $k_{22} < 1$ (resp. $k_{11} < 1$).

**Appendix 2. Model parameters and functions**

The definition and base values of parameters in our cholera models are provided in Table A1. Here $p$ (resp. $y$, $d$ and $h$) represents a person (resp. year, day and hour).

**Appendix 3. The geometric approach**

We now present the main result of the geometric approach for global stability, originally developed by Li and Muldowney [19].

We consider a dynamical system

$$\frac{dX}{dt} = f(X),$$  \hspace{1cm} (A3)

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 $(\frac{n}{2}) \times (\frac{n}{2})$ matrix-valued $C^1$ function in $D$, and set

$$U = P_J P^{-1} + PJ^{[2]} P^{-1},$$

where $P_J$ 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 $\mathcal{M}(U)$ be the Lozinski\’i measure of $U$ with respect to a matrix norm; i.e.

$$\mathcal{M}(U) = \lim_{h \to 0^+} \frac{|I + hU| - 1}{h},$$

**Table A1.** Definition of model parameters.

| Parameter | Definition | Value | References |
|-----------|------------|-------|------------|
| $N$       | Total population size of humans | 12,347 p | [21] |
| $\mu$     | Natural death rate of human | (43.5 y)$^{-1}$ | [34] |
| $\beta_l$ | Indirect transmission rate | 0.011 p$^{-1}$d$^{-1}$ | [21] |
| $\kappa$  | Half saturation rate | $10^6$ cells·ml$^{-1}$ | [17] |
| $\gamma$  | Recovery rate | $(5$ d)$^{-1}$ | [17] |
| $\sigma$  | Rate of host immunity loss | $(3$ y)$^{-1}$ | [22] |
| $\delta$  | Bacterial death rate | $(30$ d)$^{-1}$ | [17] |
| $\zeta$   | Death rate of bacteria in humans | Varied | |
| $\epsilon$ | Ratio of time scale | Varied | |
where $I$ represent the identity matrix. Define a quantity $\tilde{q}_2$ as

$$
\tilde{q}_2 = \limsup_{t \to \infty} \sup_{X_0 \in K} \frac{1}{t} \int_{0}^{t} \mathcal{M}(U(X(s,X_0))) \, ds,
$$

where $K$ is a compact absorbing subset of $D$. Then the condition $\tilde{q}_2 < 0$ 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 (65) has a unique equilibrium point $X^*$ in $D$. Then $X^*$ is globally asymptotically stable in $D$ if $\tilde{q}_2 < 0$. 