

MODELING CHOLERA TRANSMISSION UNDER DISEASE CONTROL MEASURES

MARGARET BROWN

*Department of Mathematics, Pennsylvania State University
University Park, PA 16802, USA*

MIKO JIANG

*Department of Mathematics, Brandeis University
Waltham, MA 02453, USA*

CHAYU YANG

*Department of Mathematics, University of Florida
Gainesville, FL 32611, USA*

JIN WANG*

*Department of Mathematics
University of Tennessee at Chattanooga, Chattanooga
TN 37403, USA
Jin-Wang02@utc.edu*

Received 28 October 2019

Accepted 23 May 2020

Published 11 November 2020

We present a new mathematical model to investigate the transmission dynamics of cholera under disease control measures that include education programs and water sanitation. The model incorporates the impact of education programs into the disease transmission rates and that of water sanitation into the environmental pathogen dynamics. We conduct a detailed analysis to the autonomous system of the model and establish the local and global stabilities of its equilibria that characterize the threshold dynamics of cholera. We then perform an optimal control study on the general model with time-dependent controls and explore effective approaches to implement the education programs and water sanitation while balancing their costs. Our analysis and simulation highlight the complex interaction among the direct and indirect transmission pathways of the disease, the intrinsic growth of the environmental pathogen and the impact of multiple control measures, and their roles in collectively shaping the transmission dynamics of cholera.

Keywords: Cholera; Equilibrium Analysis; Optimal Control Study.

1. Introduction

Cholera is an old disease that has been recognized since the 1600s. Today, the world is in the midst of the seventh recorded cholera pandemic, the longest one to date.^{1,2}

*Corresponding author.

The disease is caused by the bacteria *Vibrio cholerae* and is contracted by ingesting contaminated water or food. Modern technology has essentially eradicated cholera from Europe and North America, but many developing countries still experience serious outbreaks. Throughout the world, there are about 2.9 million cases and 95,000 deaths due to cholera per year according to the World Health Organization.³ Countries such as Yemen, Haiti, Zimbabwe, the DRC and areas of East Africa have seen particularly large outbreaks recently.² The reason for such cholera epidemics is complicated and involves environmental, climatic and socioeconomic factors; e.g., droughts and famines that displace populations, heavy rain that leads to increased contamination of water systems, and violent conflicts that lead to deterioration of infrastructure.^{2,4} In general, countries and regions where access to clean drinking water is not as prevalent are more likely to suffer from serious cholera outbreaks.

The recent spike in reported cholera cases and the persistence of cholera emphasizes the need for public health measures to control the disease. Common intervention methods for cholera include, among others, rehydration therapy, vaccination, and antibiotic treatment. For example, oral cholera vaccines based on killed whole cells of *Vibrio cholerae* O1 and O139 have been successfully deployed in protecting populations at high risk, and there is renewed interest in mass vaccination under outbreak and emergency settings.^{5,6} Non-pharmaceutical interventions also play an important role in reducing the disease morbidity and mortality. Water sanitation, in particular, is essential in controlling the multiplication and growth of the pathogenic bacteria in the aquatic environment and providing safe drinking water.

Meanwhile, employment of disease awareness programs is an effective approach to educate the general public on the risk and severity of cholera, offer advice on the prevention of the disease, and motivate people to make necessary changes of their routine behavior so as to reduce the possibility of infection.^{7,8} Consequently, people who are conscious of the infection risk will naturally avoid contacts with infected individuals and ingestion of contaminated water/food in order to reduce exposure to the causative pathogen.⁹ They may also attempt to adjust their routine schedules in work, travel and recreation, to pay more attention to sanitation and hygiene practice, and to receive vaccination or other preventive treatment, so as to protect themselves and their family members.^{10,11} Such disease education programs can be effectively implemented on site by health professionals, and augmented through broadcasting by television, radio stations and social media, with the common goal of communicating basic knowledge of the disease to the public and directing people toward appropriate prevention and intervention strategies.¹²

Mathematical modeling is a useful tool to quantify, and better understand, how control measures impact the transmission and spread of cholera. Several studies have recently been performed to incorporate control measures such as antibiotics, vaccination, water sanitation, and various other treatments into cholera models.^{13–16} Particularly, modeling and simulation have been conducted for the effects of awareness programs on cholera transmission.^{17,18} In these studies, however, the education programs are implemented without considering other disease control

measures, and are mostly assumed to have constant strength throughout the disease epidemic. From a practical point of view, the strength and duration of a disease awareness program should depend on the disease prevalence as well as the available resources, and would most likely vary with time. Consequently, the best way to implement an education program would be determined by the interplay between the demands and the costs of the program, and this constitutes an optimal control problem. To our knowledge, such a detailed study for cholera dynamics involving the education programs and other control measures have not been conducted thus far.

In addition, the growth of the pathogenic bacteria in the aquatic environment is an important factor in shaping the overall pattern of cholera epidemics and endemism. Earlier mathematical models assume that the bacterium *Vibrio cholerae*, the causative agent for cholera, cannot sustain itself without the contribution of human hosts.^{19,20} This assumption leads to a simple linear representation of the environmental bacterial dynamics, with a positive term representing human contribution and a negative term representing the natural removal of the bacteria. On the other hand, several cholera models^{21–23} have incorporated nonlinear intrinsic growth dynamics of the pathogenic bacteria, based on recent ecological findings that *Vibrio cholerae* indeed can independently survive and multiply in the environment,^{24,25} which leads to deeper understanding of the relationship between the environmental bacterial dynamics and cholera transmission and spread among human hosts. However, some fundamental questions such as how the intrinsic growth of the bacteria interact with cholera control measures, particularly the water sanitation, and how such interaction impacts the disease incidence and prevalence, remain unanswered at present.

Built on previous studies, this paper aims to conduct a detailed investigation on cholera transmission dynamics, with an emphasis on the interplay among the intrinsic bacterial growth, the direct and indirect transmission routes, and the multiple disease control measures including the education programs and water sanitation. To that end, we propose a new mathematical model based on differential equations. Our model incorporates both the direct, human-to-human route and the indirect, environment-to-human route that characterize the dual transmission pathways of cholera.^{4,26} We represent the impact of the education programs through the variable transmission rates, which are decreasing when the disease prevalence and bacterial concentration are increasing. Our model also includes a time-dependent function representing the impact of water sanitation. Meanwhile, the intrinsic dynamics of the bacteria are described by a logistic growth model. We conduct a rigorous analysis on the autonomous system of this model, and establish the local and global dynamics of its equilibria characterizing the threshold for disease eradication and persistence. We then perform an optimal control study on the general model with time-dependent controls and explore the most effective means to implement the education programs and water sanitation while balancing the costs of these control measures.

We organize the remainder of this paper as follows. We present the model formulation in Sec. 2, analyze the autonomous system in Sec. 3, and perform the optimal control study in Sec. 4. We conclude the paper with some discussion in Sec. 5.

2. Cholera Model

To describe the transmission dynamics of cholera under disease control measures, we consider the following system of differential equations:

$$\begin{aligned}
 \frac{dS}{dt} &= \mu N - \beta_1(I, t)SI - \beta_2(B, t)\frac{SB}{\kappa + B} - \mu S, \\
 \frac{dI}{dt} &= \beta_1(I, t)SI + \beta_2(B, t)\frac{SB}{\kappa + B} - (\gamma + \mu)I, \\
 \frac{dR}{dt} &= \gamma I - \mu R, \\
 \frac{dB}{dt} &= rB \left(1 - \frac{B}{K}\right) + \xi I - \delta(t)B,
 \end{aligned} \tag{2.1}$$

where S , I and R are the numbers of susceptible, infected and recovered individuals, respectively, and B is the concentration of the cholera bacteria in the environment. We assume that the total population, $N = S + I + R$ is a constant and μ is both the natural birth and natural death rate. All newborns enter the susceptible class. Susceptible individuals contract the disease from two transmission routes; the direct (human-to-human) and indirect (environment-to-human) transmission rates are described by the functions $\beta_1(I, t)$ and $\beta_2(B, t)$, respectively. To represent the impact of disease awareness programs, we assume that these functions are decreasing functions in I and B , respectively. This is because, as people become more educated about the risks that can lead to cholera infection, they would ideally begin to reduce contact with infected people and avoid ingesting contaminated water and food. We assume that the indirect transmission is subject to a saturating effect of the environmental bacteria with the parameter κ representing the half saturation rate. Additionally, the recovery rate from cholera is denoted by γ .

The concentration of the environmental bacteria is described by a logistic growth model with the intrinsic growth rate r and the carrying capacity K . The parameter ξ represents the shedding rate of bacteria from infected individuals back into the environment. Another control is incorporated into the model with the function $\delta(t)$, which represents the rate of removal of the bacteria due to water sanitation.

To make biological sense, we assume that the three functions $\beta_1(I, t)$, $\beta_2(B, t)$ and $\delta(t)$ are positive, bounded, and differentiable. Meanwhile, let $B = B_M$ be the positive root of $rB(1 - \frac{B}{K}) + \xi N = 0$; i.e.,

$$B_M = \frac{K + \sqrt{K^2 + 4\xi KN/r}}{2}.$$

It is then straightforward to verify that the following biologically feasible domain

$$\Omega = \{(S, I, R, B) \in \mathbf{R}_+^4 \mid S + I + R = N, B \leq B_M\}$$

is positively invariant for the vector flows of system (2.1).

3. Equilibrium Analysis

The model in Eq. (2.1) is a nonautonomous system and its analysis is difficult in general. Instead, we start our investigation by considering the simplified model where the three control functions are independent of time; i.e.,

$$\beta_1(I, t) = \beta_1(I) > 0, \quad \beta_2(B, t) = \beta_2(B) > 0, \quad \delta(t) = \delta > 0. \quad (3.1)$$

Equation (3.1) indicates that the strength of the education programs only depends on the disease prevalence and pathogen concentration, and the rate of water sanitation is a constant throughout the time. Meanwhile, since β_1 and β_2 are decreasing functions of I and B , respectively, we have

$$\beta'_1(I) \leq 0, \quad \beta'_2(B) \leq 0. \quad (3.2)$$

With the assumption (3.1), the model in (2.1) is reduced to an autonomous system and we proceed to conduct an equilibrium analysis based on the basic reproduction number, R_0 . Clearly, the disease-free equilibrium (DFE) of the system is $(N, 0, 0, 0)$. To find an expression for R_0 , we use the next-generation matrix approach in Ref. 27. This process begins with separating the equations for the infection compartments I and B into two vectors, the first one representing the rate of new infection, and the second one representing the rate of transfer into and out of the compartments. Thus, we have

$$\begin{bmatrix} \frac{dI}{dt} \\ \frac{dB}{dt} \end{bmatrix} = \hat{F} - \hat{V} = \begin{bmatrix} \beta_1(I)SI + \beta_2(B)\frac{SB}{\kappa + B} \\ rB\left(1 - \frac{B}{K}\right) + \xi I \end{bmatrix} - \begin{bmatrix} (\gamma + \mu)I \\ \delta B \end{bmatrix}.$$

From this, we obtain the Jacobians of $\hat{F}$ and $\hat{V}$ evaluated at the DFE,

$$F = \begin{bmatrix} \beta_1(0)N & \frac{\beta_2(0)N}{\kappa} \\ \xi & r \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} \gamma + \mu & 0 \\ 0 & \delta \end{bmatrix},$$

which gives us the next-generation matrix as follows:

$$FV^{-1} = \begin{bmatrix} \frac{\beta_1(0)N}{\gamma + \mu} & \frac{\beta_2(0)N}{\delta\kappa} \\ \frac{\xi}{\gamma + \mu} & \frac{r}{\delta} \end{bmatrix}.$$

Then,

$$R_0 = \rho(FV^{-1}) = \frac{1}{2} \left(\frac{\beta_1(0)N}{\gamma + \mu} + \frac{r}{\delta} + \sqrt{\left(\frac{\beta_1(0)N}{\gamma + \mu} - \frac{r}{\delta} \right)^2 + \frac{4\beta_2(0)N\xi}{\delta\kappa(\gamma + \mu)}} \right), \quad (3.3)$$

where ρ denotes the spectral radius. Equation (3.3) clearly shows that R_0 depends on both the direct and indirect transmission routes, as well as the intrinsic growth of the environmental bacteria. Specifically, let us define

$$R_h = \frac{\beta_1(0)N}{\gamma + \mu}, \quad R_e = \sqrt{\frac{\beta_2(0)N\xi}{\delta\kappa(\gamma + \mu)}}, \quad R_b = \frac{r}{\delta}, \quad (3.4)$$

where R_h and R_e represent the contribution of the direct and indirect transmission modes, respectively, to the basic reproduction number, and R_b represents the contribution of the intrinsic bacterial growth. The basic reproduction number for the autonomous system of model (2.1) is shaped by these three major components. Furthermore, it is straightforward to verify that

$$R_0 > \max(R_h, R_e, R_b). \quad (3.5)$$

This shows that if the value of any of the three components (R_h , R_e and R_b) is higher than unity, it would yield $R_0 > 1$. Biologically, this result implies that a cholera epidemic can be triggered by any (or, a combination) of the three factors: strong direct transmission, strong indirect transmission, and high bacterial growth. In particular, if we regard R_0 as a function of the intrinsic bacterial growth rate r , then direct calculation yields

$$\frac{\partial R_0}{\partial r} = \frac{1}{2\delta} \left[1 + \frac{\frac{r}{\delta} - \frac{\beta_1(0)N}{\gamma + \mu}}{\sqrt{\left(\frac{\beta_1(0)N}{\gamma + \mu} - \frac{r}{\delta} \right)^2 + \frac{4\beta_2(0)N\xi}{\delta\kappa(\gamma + \mu)}}} \right] > 0,$$

indicating that the infection risk, measured by the value of R_0 , increases monotonically with the bacterial growth rate. Our proposed disease control measures, including the education programs and water sanitation, aim to weaken the strengths of these three factors (direct and indirect transmissions, and bacterial growth) and thus reduces the risk of cholera infection.

3.1. *Disease-free equilibrium*

Based on the standard theory of the next-generation matrices and basic reproduction numbers,²⁷ we know that for $R_0 < 1$, the DFE of the system is locally asymptotically stable and for $R_0 > 1$, the DFE is unstable. We will, in fact, establish a stronger result given below.

Theorem 3.1. *The DFE of the autonomous system associated with model (2.1) is globally asymptotically stable for $R_0 \leq 1$ and unstable otherwise.*

Proof. Let $X = [I, B]^T$. We can verify that

$$X' \leq (F - V)X. \quad (3.6)$$

Indeed,

$$\begin{aligned} (F - V)X &= \left(\begin{bmatrix} \beta_1(0)N & \frac{\beta_2(0)N}{\kappa} \\ \xi & r \end{bmatrix} - \begin{bmatrix} \gamma + \mu & 0 \\ 0 & \delta \end{bmatrix} \right) \begin{bmatrix} I \\ B \end{bmatrix} \\ &= \begin{bmatrix} \beta_1(0)NI - (\gamma + \mu)I + \frac{\beta_2(0)NB}{\kappa} \\ \xi I + rB - \delta B \end{bmatrix}. \end{aligned}$$

Using the facts that $S \leq N$, $(\kappa + B) \geq \kappa$, and $rB(1 - \frac{B}{K}) \leq rB$, it follows that the inequality (3.6) holds.

Next, based on the inequality in (3.5) and some direct calculations, we can obtain that $u = [\frac{\xi}{\delta}, R_0 - \frac{\beta_1(0)N}{\gamma + \mu}]$ is a left positive eigenvector for $V^{-1}F$ corresponding to the eigenvalue of R_0 . We then construct the Lyapunov function

$$L(t) = uV^{-1}X(t).$$

From this, we have

$$L' = uV^{-1}X' \leq uV^{-1}(F - V)X = (uV^{-1}F - uV^{-1}V)X = (R_0 - 1)uX.$$

To prove the global stability, we need to pay special attention to the set of points that satisfy $L' = 0$. We can break the discussion into three cases based on the value of R_0 . Cases 1 and 2 show that the DFE is globally asymptotically stable when $R_0 \leq 1$ and Case 3 shows that the DFE is unstable when $R_0 > 1$.

Case 1. $R_0 < 1$. In this case, $L' \leq 0$. To satisfy $L' = 0$, it must be true that $uX = 0$, thus $X = 0$ since u is positive. That means $I = B = 0$, which implies that the largest invariance set for $L' = 0$ is the singleton $(N, 0, 0, 0)$; i.e., the DFE. Hence, based on LaSalle's invariance principle, the DFE is globally asymptotically stable when $R_0 < 1$.

Case 2: $R_0 = 1$. In this case, by equating our expression of R_0 to 1, we can do some algebra to obtain

$$\beta_2(0)\xi = \frac{(\gamma + \mu - \beta_1(0)N)(\delta - r)\kappa}{N}, \quad (3.7)$$

where $\gamma + \mu - \beta_1(0)N > 0$ and $\delta - r > 0$ since $R_0 = 1 > \max\{\frac{\beta_1(0)N}{\gamma + \mu}, \frac{r}{\delta}\}$. Then,

$$L' = 0 \Rightarrow uV^{-1}X' = 0 \Rightarrow \frac{dI}{dt}\xi + \frac{dB}{dt}(\gamma + \mu - \beta_1(0)N) = 0.$$

Substituting the equations for $\frac{dI}{dt}$ and $\frac{dB}{dt}$ from our system (2.1), we obtain

$$\xi\beta_1(0)I(S - N) + \xi\beta_2(0)\frac{SB}{\kappa + B} + (\gamma + \mu - \beta_1(0)N)\left(r - \delta - \frac{rB}{K}\right)B = 0. \quad (3.8)$$

Substituting (3.7) into (3.8) yields

$$\xi\beta_1(0)I(S - N) + B(\gamma + \mu - \beta_1(0)N)\left((\delta - r)\frac{\kappa(S - N) - BN}{N(\kappa + B)} - \frac{rB}{K}\right) = 0. \quad (3.9)$$

Since $S \leq N$, we must have $B = 0$ and either $S = N$ or $I = 0$ for equality (3.9) to hold. Either way, this means that the largest invariance set of $L' = 0$ is the singleton $(N, 0, 0, 0)$. Again, LaSalle's invariance principle yields that the DFE is globally asymptotically stable when $R_0 = 1$.

Case 3. $R_0 > 1$. In this case, by continuity, L' will also be greater than 0 in some neighborhood of the DFE. Then, by the Lyapunov Stability theorem, the DFE will be unstable. $\square$

3.2. Endemic equilibrium

Next, we will turn our attention to a nontrivial equilibrium point. Such a point would satisfy

$$\mu N = \beta_1(I)SI + \beta_2(B)\frac{SB}{\kappa + B} + \mu S, \quad (3.10)$$

$$(\gamma + \mu)I = \beta_1(I)SI + \beta_2(B)\frac{SB}{\kappa + B}, \quad (3.11)$$

$$0 = \gamma I - \mu R, \quad (3.12)$$

$$0 = rB\left(1 - \frac{B}{K}\right) + \xi I - \delta B. \quad (3.13)$$

From Eq. (3.13), we can write I as a function of B as follows:

$$I(B) = \frac{r}{\xi K}B^2 + \frac{\delta - r}{\xi}B. \quad (3.14)$$

Note that for our system to have biological meaning, we only consider the case when I and B are positive. We will consider two cases: when $\delta > r$ and when $\delta \leq r$.

First, when $\delta > r$, the function $I(B)$ is an increasing function for $B > 0$. We can combine Eqs. (3.10) and (3.11) into one equation by substituting $(\gamma + \mu)I$ for $\beta_1(I)SI + \beta_2(B)\frac{SB}{\kappa + B}$. This yields

$$S(B) = N - \frac{\gamma + \mu}{\mu}I(B). \quad (3.15)$$

Since $I(B)$ is an increasing function, $S(B)$ is a decreasing function. Next, we rewrite Eq. (3.11) so that everything is a function of B and obtain

$$(\gamma + \mu)I(B) = \beta_1(I(B))S(B)I(B) + \beta_2(B)\frac{S(B)B}{\kappa + B}. \quad (3.16)$$

Substituting Eq. (3.14) to Eq. (3.16), we can cancel B from both sides and rearrange to obtain

$$\frac{\beta_2(B)S(B)}{(\kappa + B)\left(\frac{r}{\xi K}B + \frac{\delta - r}{\xi}\right)} + \beta_1(I(B))S(B) = \gamma + \mu. \quad (3.17)$$

Now we determine whether there is a unique positive B that is a solution to this equation. Denote the left-hand side of Eq. (3.17) by $f(B)$. Then, $f(B)$ is a decreasing function, because we know $\beta_1(B)$, $\beta_2(B)$, and $S(B)$ are decreasing and the denominator is increasing and positive due to the assumption $\delta > r$. Therefore, in order to guarantee a unique solution, we must have $f(0) > \gamma + \mu$; i.e.,

$$\frac{\beta_2(0)\xi N}{\kappa(\delta - r)} + \beta_1(0)N > \gamma + \mu,$$

which is equivalent to the condition $R_0 > 1$. To see this, we use our previous expression for R_0 and manipulate the algebra as follows:

$$\begin{aligned} R_0 > 1 &\Leftrightarrow \left(\frac{\beta_1(0)N}{\gamma + \mu} - \frac{r}{\delta}\right)^2 + \frac{4\beta_2(0)N\xi}{\delta\kappa(\gamma + \mu)} > \left(2 - \frac{\beta_1(0)N}{\gamma + \mu} - \frac{r}{\delta}\right)^2 \\ &\Leftrightarrow \frac{\beta_2(0)N\xi}{\delta\kappa(\gamma + \mu)} > 1 - \frac{\beta_1(0)N}{\gamma + \mu} - \frac{r}{\delta} + \frac{\beta_1(0)Nr}{\delta(\gamma + \mu)} \\ &\Leftrightarrow \beta_2(0)\xi N > \delta\kappa(\gamma + \mu) - \delta\kappa\beta_1(0)N - r\kappa(\gamma + \mu) + \beta_1(0)Nr\kappa \\ &\Leftrightarrow \beta_2(0)\xi N > \kappa(\delta - r)(\gamma + \mu - \beta_1(0)N) \\ &\Leftrightarrow \frac{\beta_2(0)\xi N}{\kappa(\delta - r)} + \beta_1(0)N > \gamma + \mu. \end{aligned}$$

Next we consider the case when $\delta \leq r$. In this case, since we only consider when I and B are greater than zero, $I(B)$ will still be increasing for $B > B_1$ where B_1 is the positive horizontal intercept of $I(B)$. Then, the calculations from the previous case will still hold. Thus, Eq. (3.17) and the observation that $f(B)$ is decreasing are still valid.

Observe that as B approaches B_1 from the right, $f(B)$ goes to infinity because the denominator $(\kappa + B)\left(\frac{r}{\xi K}B + \frac{\delta - r}{\xi}\right)$ approaches 0. On the other hand, based on Eq. (3.15), there is a $B_2 > B_1$ such that $S(B) \leq 0$ for $B \geq B_2$, which indicates that $f(B) \leq 0$ for $B \geq B_2$. Thus, $\delta \leq r$ implies that there is a unique nontrivial equilibrium $B^* \in (B_1, B_2)$. Lastly, note that $\delta \leq r$ implies $R_0 > \frac{r}{\delta} \geq 1$.

Hence, the existence and uniqueness of the endemic equilibrium can be summarized in the theorem below.

Theorem 3.2. *The autonomous system associated with model (2.1) has a unique endemic equilibrium if and only if $R_0 > 1$.*

We have conducted detailed stability analysis on this endemic equilibrium. Regarding its local stability, we have the following theorem and the proof is provided in Appendix A.

Theorem 3.3. *When $R_0 > 1$, the endemic equilibrium of the autonomous system associated with model (2.1) is locally asymptotically stable provided that condition (A.5) holds.*

We have also analyzed the global asymptotical stability of the endemic equilibrium. We state the theorem below, with the proof given in Appendix B.

Theorem 3.4. *When $R_0 > 1$, the endemic equilibrium of the autonomous system associated with model (2.1) is globally asymptotically stable in the interior of the domain Ω if $\beta_1(0) \leq \frac{\gamma}{2N}$.*

Note that $\beta_1(I)$ is assumed to be a decreasing function of I . Thus, the sufficient condition $\beta_1(0) \leq \frac{\gamma}{2N}$ in Theorem 3.4 provides an upper bound for the direction transmission rate β_1 to ensure the global asymptotical stability of the endemic equilibrium.

4. Optimal Control

Through a detailed equilibrium analysis, we have resolved the main dynamics of the simplified, autonomous system of our cholera model (2.1). Particularly, Theorems 3.1 and 3.4 establish $R_0 = 1$ as a sharp threshold for disease eradication and disease persistence. Now we consider the original, nonautonomous system (2.1) with time-dependent functions $\beta_1(I, t)$, $\beta_2(B, t)$ and $\delta(t)$. These three functions are directly related to the human-to-human transmission, the environment-to-human transmission, and the environmental bacterial growth, respectively. Per our discussion in the previous section, the infection risk (measured by the basic reproduction number R_0) of the autonomous system is shaped by these three factors. Since the autonomous system can be regarded as an approximation, in the time-average sense, of the time-dependent system (2.1), it is natural to expect that these three components (direct and indirect transmissions, and bacterial growth) also play a critical role for system (2.1) and that a reduction in the strength of these components, represented by the three functions $\beta_1(I, t)$, $\beta_2(B, t)$ and $\delta(t)$, respectively, would reduce the infection risk.

As stated before, $\beta_1(I, t)$ and $\beta_2(B, t)$ represent the impact of education programs, whereas $\delta(t)$ describes the effect of water sanitation. We are interested in

knowing how the changes of these functions with respect to time would impact the disease transmission and spread. Meanwhile, since each of these disease control measures comes with a cost, we are also interested in seeing how the costs of the control are factored into this process. Thus, the goal of our study on the nonautonomous system (2.1) is to explore a way that the effects and the costs of the disease control measures can best be balanced based on the functions $\beta_1(I, t)$, $\beta_2(B, t)$ and $\delta(t)$. We achieve this goal through an optimal control study.

There are many ways to design biologically meaningful, time-dependent disease control functions. In our model formulation, the direct transmission rate $\beta_1(I, t)$ is a decreasing function of I , and the strength of the incorporated control measure would naturally be inversely correlated to that of the direct transmission. Similarly, the indirect transmission rate $\beta_2(B, t)$ is a decreasing function of B , and the strength of the respective control measure would naturally be negatively correlated to that of the indirect transmission. For a simple choice to reflect these properties, we set the two transmission functions as

$$\beta_1(I, t) = a_1 - b_1(t)I/N \quad \text{and} \quad \beta_2(B, t) = a_2 - b_2(t)B/K,$$

where a_1 and a_2 are positive constants representing the upper bounds of the transmission rates and $b_1(t)$ and $b_2(t)$ are time-dependent, nonnegative functions that specify the impact of the education programs in reducing the disease transmission rates. We study the system on a time interval $[0, T]$ with the control set

$$\Gamma = \{(\delta(t), b_1(t), b_2(t)) \mid 0 \leq \delta(t) \leq \delta_{\max}, 0 \leq b_1(t) \leq a_1, 0 \leq b_2(t) \leq a_2\}. \quad (4.1)$$

Our goal is to minimize the total number of infections and the costs of the controls over the time period $[0, T]$; i.e.,

$$\min_{(b_1, b_2, \delta) \in \Gamma} \int_0^T [I(t) + c_1 b_1(t)^2 + c_2 b_2(t)^2 + c_3 \delta(t)^2] dt, \quad (4.2)$$

where the positive constants c_i , $1 \leq i \leq 3$, are the costs parameters for the respective education programs and water sanitation. These cost parameters serve as balancing coefficients that transform the integral of the three control terms into the amount of dollars in a practical sense. We have used quadratic terms to represent the costs, which is common in optimal control studies.^{14,28,29}

For the control set Γ , we note that it is closed and connected and the integrand of (4.2) is convex. We also note that our model is linear in the control variables. Based on the standard optimal control theorems,^{30,31} we obtain the following result.

Theorem 4.1. *There exists $(\delta^*(t), b_1^*(t), b_2^*(t)) \in \Gamma$ such that the objective functional in (4.2) is minimized.*

Furthermore, the uniqueness of the optimal control solution is guaranteed for small time period T due to the Lipschitz property of the model equations and

the boundedness of the state variables.³⁰ To determine the optimal control solution, Pontryagin's Maximum/Minimum Principle³¹ provides a useful tool for constructing the optimal control system which is then numerically solved. This procedure basically transfers the problem of minimizing (4.2) into one of minimizing the Hamiltonian with respect to the controls.

As before, the fact that the total population size N remains constant allows us to eliminate the equation for R . These equations of our original system (2.1) are referred to as the state equations. We define the corresponding adjoint functions as λ_S , λ_I , and λ_B . The Hamiltonian $\mathcal{H}$ is constructed by multiplying each adjoint function with the right-hand side of its corresponding state equation, and adding the integrand of the objective functional to the sum of these products. In this case, we have

$$\begin{aligned} \mathcal{H} = & I(t) + c_1 b_1(t)^2 + c_2 b_2(t)^2 + c_3 \delta(t)^2 \\ & + \lambda_S \left(\mu N - \left(\beta_2(B, t) \frac{B}{\kappa + B} + \beta_1(I, t) I \right) S - \mu S \right) \\ & + \lambda_I \left(\left(\beta_2(B, t) \frac{B}{\kappa + B} + \beta_1(I, t) I \right) S - (\gamma + \mu) I \right) \\ & + \lambda_B \left(rB \left(1 - \frac{B}{K} \right) + \xi I - \delta(t) B \right). \end{aligned} \quad (4.3)$$

To achieve the optimal control, the adjoint functions must satisfy

$$\frac{d\lambda_S}{dt} = -\frac{d\mathcal{H}}{dS}, \quad \frac{d\lambda_I}{dt} = -\frac{d\mathcal{H}}{dI}, \quad \frac{d\lambda_B}{dt} = -\frac{d\mathcal{H}}{dB}.$$

That is

$$\frac{d\lambda_S}{dt} = (\lambda_S - \lambda_I) \left(\frac{a_2 B}{\kappa + B} - \frac{b_2(t) B^2}{K(\kappa + B)} + a_1 I - b_1(t) \frac{I^2}{N} \right) + \mu \lambda_S, \quad (4.4)$$

$$\frac{d\lambda_I}{dt} = -1 + (\lambda_S - \lambda_I) \left(a_1 S - 2b_1(t) \frac{IS}{N} \right) + \lambda_I(\gamma + \mu) - \lambda_B \xi, \quad (4.5)$$

$$\frac{d\lambda_B}{dt} = (\lambda_S - \lambda_I) \frac{a_2 \kappa K S - b_2(t) S B (2\kappa + B)}{K(\kappa + B)^2} + \lambda_B \left(\frac{2rB}{K} + \delta(t) - r \right), \quad (4.6)$$

with the final time conditions: $\lambda_S(T) = 0$, $\lambda_I(T) = 0$, and $\lambda_B(T) = 0$. Meanwhile, the optimal control is characterized by

$$\delta^*(t) = \max [0, \min(\tilde{\delta}(t), \delta_{\max})], \quad (4.7)$$

$$b_1^*(t) = \max [0, \min(\tilde{b}_1(t), a_1)], \quad (4.8)$$

$$b_2^*(t) = \max [0, \min(\tilde{b}_2(t), a_2)], \quad (4.9)$$

where $\tilde{\delta}(t) = \frac{\lambda_B B}{2c_3}$, $\tilde{b}_1(t) = \frac{(\lambda_I - \lambda_S)I^2 S}{2c_1 N}$, and $\tilde{b}_2(t) = \frac{(\lambda_I - \lambda_S)B^2 S}{2c_2(\kappa + B)K}$ come from the conditions

$$\frac{\partial \mathcal{H}}{\partial \delta} = 0, \quad \frac{\partial \mathcal{H}}{\partial b_1} = 0, \quad \frac{\partial \mathcal{H}}{\partial b_2} = 0,$$

respectively.

In short, this optimal control model consists of the state system (2.1) with its initial conditions, the adjoint Eqs. (4.4)–(4.6) with the final conditions, and the optimal controls characterized by (4.7)–(4.9). A numerical method is necessary to solve these coupled equations. Since the state system has initial conditions, but the adjoint system has final conditions, it is not possible to numerically solve the entire system all at once. Instead, such a problem can be solved by applying the forward-backward sweep method.²⁸ This consists of solving the state system forward in time, then solving the adjoint system backward in time, and then updating the control functions. Each step uses the information gained from the preceding step, and the process is repeated until it converges sufficiently. In our case, we use a fourth-order Runge–Kutta method to compute the state and adjoint equations in each iteration.

The parameter values used in our simulation are listed in Table 1.

Figure 1 compares the infected population under two cases: with the optimal control and without any control. Here, the number of infected individuals is relative to the (normalized) total population $N = 10,000$. We assume that the costs for all the three control measures are low and comparable to each other; for simplicity, we set the cost parameters as $c_1 = c_2 = c_3 = 1$. It is clear from the graph that when the optimal control is implemented, the peak value of the infected population is significantly reduced (approximately 30% reduction), and the number of infected individuals eventually approaches zero, in contrast to the case without any control where the infection persists at a positive level. Meanwhile, Fig. 1 also shows the bacterial concentration with and without the optimal control. We see that when there is no control, the bacterial concentration quickly increases and then stabilizes at a high level. In contrast, when the optimal control is incorporated, the bacterial

Table 1.

Parameter	Symbol	Value	Unit	Source
Normalized total population	N	10,000	person	—
Natural human birth/death rate	μ	6.3×10^{-5}	1/day	Ref. 13
Recovery rate	γ	0.05	1/day	Ref. 13
Half-saturation rate of bacteria	κ	10^6	cell/ml	Ref. 14
Carrying capacity of bacteria	K	10^8	cell/ml	Ref. 14
Bacterial intrinsic growth rate	r	0.1	1/day	—
Human shedding rate	ξ	20	cell/ml/day	Ref. 13
Cost parameter for education program b_1	c_1	varied	dollar · day · person ²	—
Cost parameter for education program b_2	c_2	varied	dollar · day	—
Cost parameter for water sanitation δ	c_3	varied	dollar · day	—
Upper bound of direct transmission rate	a_1	0.05	1/person/day	Ref. 13
Upper bound of indirect transmission rate	a_2	0.2	1/day	Ref. 13
Upper bound of water sanitation rate	$\delta_{\max}$	0.3	1/day	—

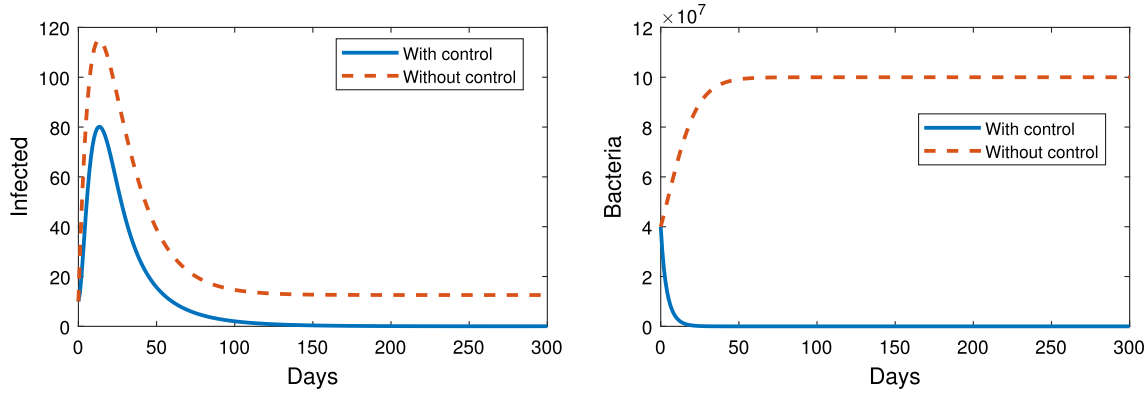


Fig. 1. Optimal control results with $c_1 = c_2 = c_3 = 1$. Left: the infected population with and without the optimal control; Right: the bacteria concentration with and without the optimal control.

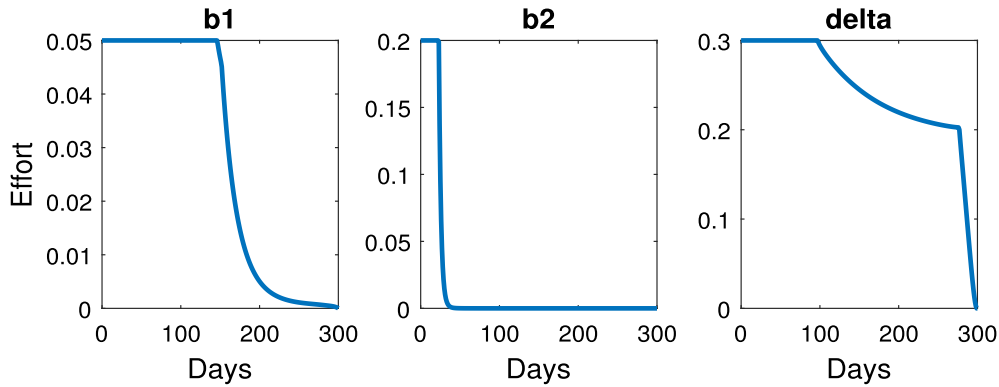


Fig. 2. Optimal control profiles with $c_1 = c_2 = c_3 = 1$.

concentration sharply decreases to, and then remains at, a level very close to zero, consistent with the observation that the disease prevalence approaches zero in the long run.

The profiles of the optimal control functions are displayed in Fig. 2. We observe that each optimal control starts at its maximum and stays at that level for a certain period of time, a common pattern that has been observed in many optimal control studies^{13–15,29}; afterwards it starts decreasing and goes to zero eventually. For the profile of $b_1(t)$, which is related to the human-to-human transmission, it stays at the maximum level for about 100 days before decreasing, and this turning point seems to be correlated to the time that the number of infected individuals is approaching zero (see Fig. 1). For the profile of $b_2(t)$, which is related to the environment-to-human transmission, it stays at its maximum for a much shorter period (about 25 days) before decreasing to zero, possibly due to the fact that the environmental bacterial concentration quickly approaches a level very close to zero during the first 20–25 days (see Fig. 1). Additionally, for the profile of $\delta(t)$, which is related to water sanitation, there is a significant transient period between the time it stays

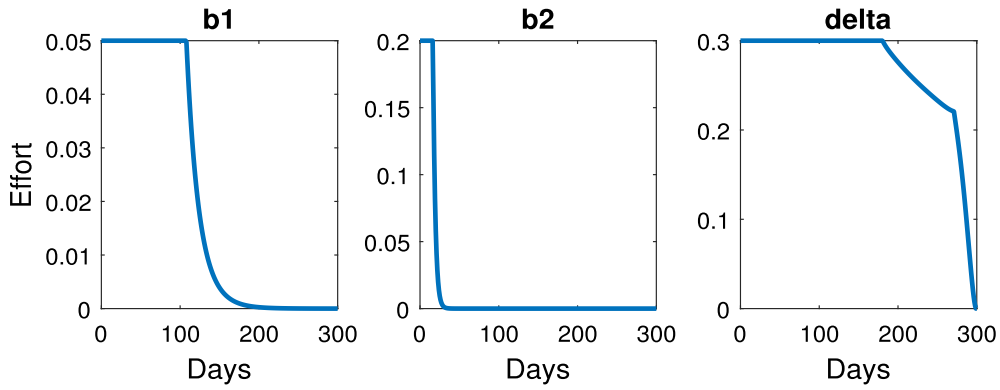


Fig. 3. Optimal control profiles with $c_1 = c_2 = 10$ and $c_3 = 0.1$.

at the maximum and the time it decreases to zero. This transient period, where the strength of water sanitation remains at a relatively high level, is needed to ensure that the bacterial concentration stays at a very low level so that the disease transmission risk is minimized.

It is natural to expect that when the costs of these control methods are changed, the optimal control profiles will also change. Figure 3 shows the control profiles where the costs of the education programs are increased by 10 times (compared to the previous setting) while the cost of water sanitation is decreased by 10 times; i.e., $c_1 = c_2 = 10$ and $c_3 = 0.1$. Compared to Fig. 2, we see that the lengths of time for the maximum strength of the two education programs are both shortened, due to the increased costs, whereas the period for the maximum strength of water sanitation is extended because of the decreased cost. An opposite pattern is displayed in Fig. 4 where the costs of the education programs are decreased by 10 times and the cost of water sanitation is increased by 10 times; i.e., $c_1 = c_2 = 0.1$ and $c_3 = 10$. In both figures, we see that the changes for the $b_2(t)$ profile are moderate, whereas the $b_1(t)$ and $\delta(t)$ profiles are more sensitive to the costs and exhibit more significant changes.

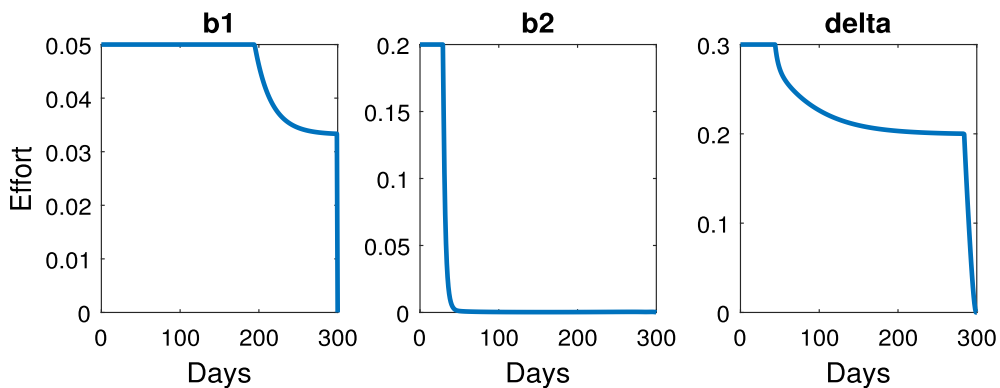


Fig. 4. Optimal control profiles with $c_1 = c_2 = 0.1$ and $c_3 = 10$.

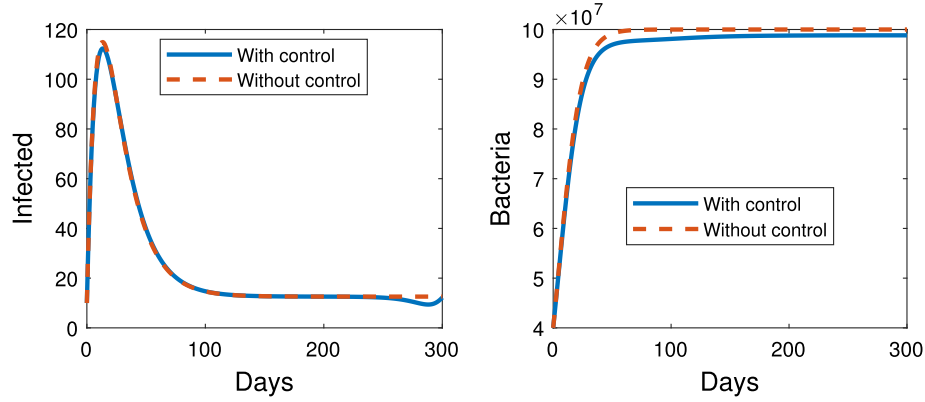


Fig. 5. Optimal control results with $c_1 = 100$, $c_2 = 500$ and $c_3 = 1000$. Left: infection population with and without the optimal control; Right: bacteria population with and without the optimal control.

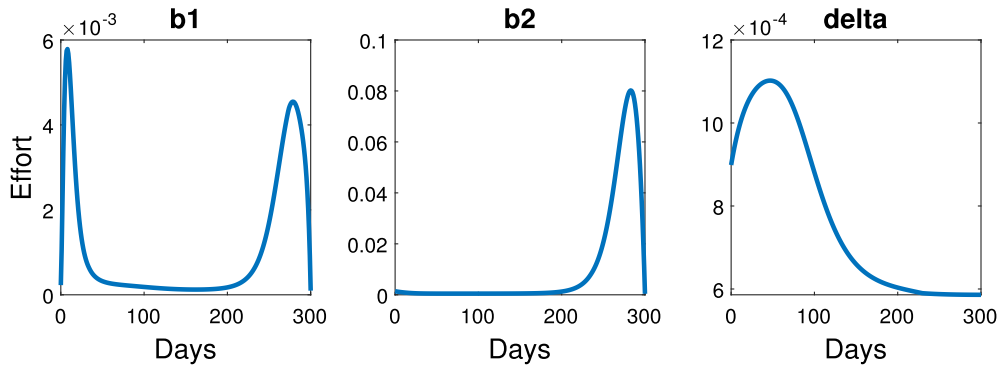


Fig. 6. Optimal control profiles with $c_1 = 100$, $c_2 = 500$ and $c_3 = 1000$.

In addition, if the costs of these control measures are prohibitively high, then the disease control would have little impact on the reduction of the infected population and the bacterial concentration. Figures 5 and 6 illustrate such a scenario where $c_1 = 100$, $c_2 = 500$ and $c_3 = 1000$. In this case, the optimal control solution tries to adjust to the high costs by significantly reducing the strength and duration of the control measures, leading to little or no improvement in terms of disease prevalence and pathogen burden.

Finally, we have also tested the scenarios of applying one individual control at a time. To do that, each time we set two of the three control variables, $b_1(t)$, $b_2(t)$ and $\delta(t)$, to be identically zero so that only one control is effective in the model. Then, we follow a similar procedure of optimal control analysis and simulation for that single control and output the results. We find that generally the outcome of a single optimal control is not as good as that from the optimal combination of the three controls. For example, Fig. 7 illustrates a case where only the direct transmission control $b_1(t)$ is implemented, with a low cost $c_1 = 1$, and $b_2(t) = \delta(t) \equiv 0$. We see that the peak value of the prevalence is reduced by approximately

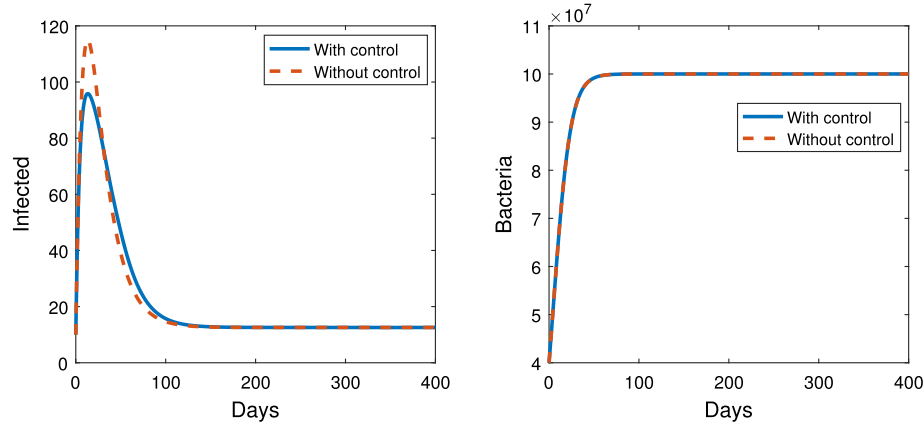


Fig. 7. Optimal control results with one single control $b_1(t)$ on the education program, where $c_1 = 1$. Left: the infected population with and without the optimal control; Right: the bacteria concentration with and without the optimal control.

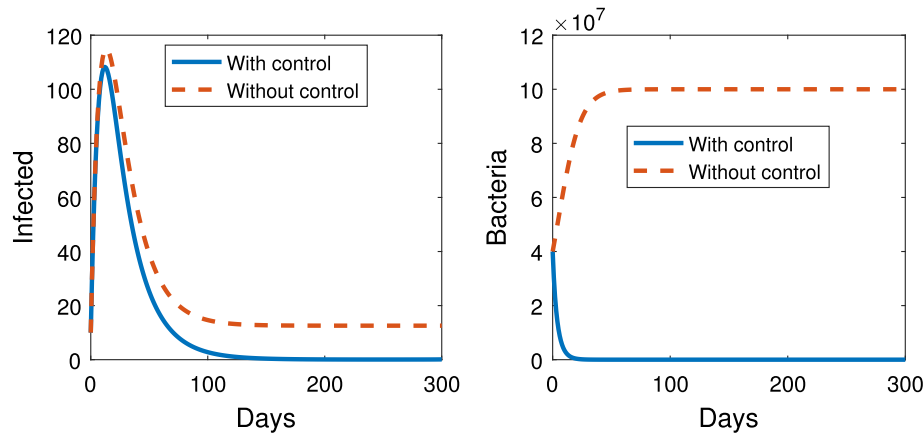


Fig. 8. Optimal control results with one single control $\delta(t)$ on the water sanitation, where $c_3 = 1$. Left: the infected population with and without the optimal control; Right: the bacteria concentration with and without the optimal control.

15% with the optimal control, though the infection persists at the endemic level in the long run. The bacterial concentration, as can be naturally expected, shows almost no difference with or without this single control. Compared to Fig. 1, it is clear that the impact of this single control is not as significant as that of the three controls combined. Meanwhile, Fig. 8 shows a typical case where only the water sanitation control $\delta(t)$ is implemented, with a low cost $c_3 = 1$, and $b_1(t) = b_2(t) \equiv 0$. We observe that the bacterial concentration quickly decays to 0, a pattern similar to that in Fig. 1, but there is only a very minor improvement in reducing the infection peak under this control. These results demonstrate that each individual control targets a specific part of cholera transmission dynamics and that a strategic combination of these three control measures generally achieves the best outcome.

5. Conclusion

Cholera has become a threat to populations in developing countries, highlighting the need for disease control or even eradication. In this paper, we have presented a new mathematical model to investigate the transmission dynamics of cholera under disease control measures that include education programs and water sanitation. The inclusion of such controls provides a two-pronged approach, where the education programs target the direct and indirect transmission rates while water sanitation reduces the population growth of the environmental bacteria. We have conducted a detailed equilibrium analysis and optimal control study on this model, and our results indicate that water sanitation and public awareness programs are both feasible intervention methods for reducing disease prevalence and controlling a cholera epidemic.

Our study has incorporated the multiple transmission pathways (both the direct and indirect routes), the multiple disease intervention methods (both the education programs and water sanitation), and the environmental bacterial dynamics, into a single cholera modeling framework. A focus of this paper is to explore the complex interactions among the environment-to-human transmission route, the human-to-human transmission route, and the intrinsic growth of the environmental pathogen, and to understand how the disease awareness programs and water sanitation impact these three critical components of cholera epidemics and reduce the transmission and spread of cholera. In particular, our equilibrium analysis of the autonomous system shows that the disease transmission risk, measured by the basic reproduction number, is shaped by the interplay of the dual transmission pathways and the intrinsic bacterial dynamics.

We have put an emphasis on the optimal control study of our cholera control measures. Optimal control study has become an important part of mathematical and computational epidemiology that can provide useful guidelines for designing effective disease intervention strategies while balancing the costs of the control measures. There are, however, relatively few studies devoted to the optimal control of cholera, and no results have been published yet for the optimal implementation of the education programs and optimal combination of those programs and other intervention methods such as water sanitation. This paper partially fills this knowledge gap. Our simulation results show that the education programs and water sanitation can effectively reduce the disease prevalence and environmental bacterial concentration, that these control measures should start at their maximum strength whenever resources allow in order to best control the disease, and that the cost factors play an important role in determining the optimal control profiles. Our results also demonstrate that an optimal combination of these multiple control measures generally achieves a better outcome than using a single control method. These findings could provide useful guidelines to public health administrations in the design of more effective prevention and intervention strategies for cholera.

Overall, the analysis and simulation presented in this paper lead the way to a deeper understanding of cholera transmission. Future work along this line could incorporate other epidemiological, environmental and socioeconomic factors, such as the age structure of the host population, the seasonal fluctuation of the disease incidence, and the spatial heterogeneity of the pathogen and host settings, for an improved modeling framework and for a more holistic study of cholera dynamics.

Acknowledgments

We acknowledge partial support from the NSF REU program under grant number 1852288 which provided an opportunity for the first two authors (M.B. and M.J.) to conduct undergraduate research at UTC. We are grateful to the managing editor and the three anonymous referees for their helpful comments that have improved the original paper.

References

1. Ali M, Nelson AR, Lopez AL, Sack DA, Updated Global Burden of Cholera in Endemic Countries, *PLoS Negl Trop Dis* **9**(6):e0003832, 2015.
2. World Health Organization: Weekly Epidemiological Report, 21 September 2018, **93**(38):489–500, 2018, <http://www.who.int/wer/2018/wer9338/en/>.
3. World Health Organization: Ending Cholera: A Global Roadmap to 2030, 3 October 2017, <https://www.who.int/cholera/publications/global-roadmap.pdf?ua=1>.
4. Nelson EJ, Harris JB, Morris JG, Calderwood SB, Camilli A, Cholera transmission: The host, pathogen and bacteriophage dynamics, *Nat Rev Microbiol* **7**:693–702, 2009.
5. Longini IM, Nizam A, Ali M, Yunus M, Shenvi N, Clemens JD, Controlling endemic cholera with oral vaccines, *PLoS Med* **4**(11):e336, 2007.
6. Lucas ME, Deen JL, von Seidlein L *et al.*, Effectiveness of mass oral cholera vaccination in Beira, Mozambique, *N Engl J Med* **352**(8):757–767, 2005.
7. Funk S, Gilad E, Janse VAA, Endemic disease, awareness, and local behavioural response, *J Theor Biol* **264**:501–509, 2010.
8. Funk S, Salathé M, Jansen VAA, Modelling the influence of human behaviour on the spread of infectious diseases: A review, *J. R. Soc. Interface* **7**(50):1247–1256, 2010.
9. Einarsdóttir J, Passa A, Gunnlaugsson G, Health education and cholera in rural Guinea-Bissau, *Int J Infect Dis* **5**(3):133–138, 2001.
10. Philipson T, Private vaccination and public health: An empirical examination for US measles, *J Hum Resour* **31**:611–630, 1996.
11. Leung GM, Lam TH, Ho LM, Ho SY, Chan BH, Wong IO, and Hedley AJ, The impact of community psychological responses on outbreak control for severe acute respiratory syndrome in Hong Kong, *J Epidemiol Commun Heal* **57**(11):857–863, 2003.
12. World Health Organization: Public health campaigns: Getting the message across, Geneva, Switzerland, 2009.
13. Posny D, Wang J, Mukandavire Z, Modnak C, Analyzing transmission dynamics of cholera with public health interventions, *Math Biosci* **264**:38–53, 2015.
14. Neilan RLM, Schaefer E, Gaff H, Fister KR, Lenhart S, Modeling Optimal Intervention Strategies for Cholera, *Bull Math Biol* **72**:2004–2018, 2010.

15. Cai LM, Modnak C, Wang J, An age-structured model for cholera control with vaccination, *Appl Math Comput* **299**:127–140, 2017.
16. Yang J, Modnak C, and Wang J, Dynamical analysis and optimal control simulation for an age-structured cholera transmission model, *J Franklin Inst* **356**:8438–8467, 2019.
17. Wang X, Gao D, Wang J, Influence of human behavior on cholera dynamics, *Math Biosci* **267**:41–52, 2015.
18. Yang C, Wang X, Gao D, Wang J, Impact of awareness programs on cholera dynamics: Two modeling approaches, *Bull Math Biol* **79**(9):2109–2131, 2017.
19. Codeço CT, Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir, *BMC Infect Dis* **1**:1, 2001.
20. Hartley DM, Morris JG, Smith DL, Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics?, *PLoS Med.* **3**:0063–0069, 2006.
21. Yang C, Wang J, On the intrinsic dynamics of waterborne infections, *Math Biosci* **296**:71–81, 2018.
22. Wang X, Wang J, Analysis of cholera epidemics with bacterial growth and spatial movement, *J Biol Dyn* **9**(1):233–261, 2015.
23. Posny D, Wang J, Modelling cholera in periodic environments, *J Biol Dyn* **8**(1):1–19, 2014.
24. Colwell RR, A global and historical perspective of the genus *Vibrio*, in Thompson FL, Austin B, Swings J (eds.) *The Biology of Vibrios* ASM Press, Washington DC, 2006.
25. Faruque SM, Nair GB, *Vibrio Cholerae: Genomics and Molecular Biology*, Caister Academic Press, 2008.
26. Mukandavire Z, Liao S, Wang J, Gaff H, Smith DL, Morris JG, Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe, *Proc Natl Acad Sci USA* **108**:8767–8772, 2011.
27. van den Driessche P, Watmough J, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math Biosci* **180**:29–48, 2001.
28. Lenhart S, Workman J, *Optimal Control Applied to Biological Models*, Chapman Hall/CRC, 2007.
29. Jung E, Iwami S, Takeuchi Y, Jo T-C, Optimal control strategy for prevention of avian influenza pandemic, *J. Theor. Biol.* **260**:220–229, 2009.
30. Fleming WH, Rishel RW, *Deterministic and Stochastic Optimal Control*, Springer, New York, 1975.
31. Pontryagin LS, Boltyanski VG, Gamkrelize RV, Mishchenko EF, *The mathematical Theory of Optimal Processes*, Wiley, New York, 1967.
32. Li MY, Muldowney JS, A geometric approach to global-stability problems, *SIAM J Math Anal* **27**:1070–1083, 1996.
33. Muldowney JS, Compound matrices and ordinary differential equations, *Rocky M J Math* **20**:857–872, 1990.

Appendix A. Proof of Theorem 3.3

We use the Routh–Hurwitz criteria to show that the endemic equilibrium of the autonomous system is locally asymptotically stable. The criteria provide simple algebraic conditions that are necessary and sufficient for the stability of polynomials. To apply this, start with the Jacobian of the system, evaluated at the endemic

equilibrium $X^* = (S^*, I^*, R^*, B^*)$,

$$J = \begin{bmatrix} -\beta_1(I^*)I^* - \frac{\beta_2(B^*)B^*}{\kappa + B^*} - \mu & -a & 0 & -b \\ \beta_1(I^*)I^* + \frac{\beta_2(B^*)B^*}{\kappa + B^*} & a - (\gamma + \mu) & 0 & b \\ 0 & \gamma & -\mu & 0 \\ 0 & \xi & 0 & r - \delta - \frac{2rB^*}{K} \end{bmatrix},$$

where

$$a = S^*\beta_1(I^*) + S^*I\beta'_1(I^*), \quad b = \frac{\beta_2(B^*)S^*\kappa}{(\kappa + B^*)^2} + \frac{\beta'_2(B^*)S^*B^*}{(\kappa + B^*)}.$$

Then, the characteristic polynomial is $\det(\lambda I - J) =$

$$\begin{aligned} & (\lambda + \mu) \det \begin{bmatrix} -\beta_1(I^*)I^* - \frac{\beta_2(B^*)B^*}{\kappa + B^*} - \mu & -a & -b \\ \beta_1(I^*)I^* + \frac{\beta_2(B^*)B^*}{\kappa + B^*} & a - (\gamma + \mu) & b \\ 0 & \xi & r - \delta - \frac{2rB^*}{K} \end{bmatrix} \\ & = (\lambda + \mu)(\lambda^3 + x\lambda^2 + y\lambda + z), \end{aligned}$$

where for ease of presentation, we write the entries of the above 3×3 matrix in a general form

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

It follows that the value of x, y, z are

$$x = -(a_{11} + a_{22} + a_{33}), \quad (\text{A.1})$$

$$y = a_{22}a_{33} - \xi a_{23} - a_{12}a_{21} + a_{11}(a_{22} + a_{33}), \quad (\text{A.2})$$

$$z = \xi(a_{11}a_{23} - a_{13}a_{21}) - a_{11}a_{22}a_{33} + a_{12}a_{21}a_{33}. \quad (\text{A.3})$$

To satisfy the Routh–Hurwitz criteria for stability, we show that

$$x > 0, \quad y > 0, \quad z > 0, \quad \text{and} \quad xy > z.$$

Note that from Eq. (3.11), we have $S\beta_1(I^*) - (\gamma + \mu) = -\beta_2(B^*)\frac{S^*B^*}{I^*(\kappa + B^*)}$. Then

$$a_{22} = -\beta_2(B^*)\frac{S^*B^*}{\kappa + B^*} + S^*I^*\beta'_1(I^*) < 0,$$

since $\beta_2(B^*) > 0$ and $\beta'_1(I^*) \leq 0$. Similarly, based on Eq. (3.13), we can rewrite $r - \delta$ to obtain

$$a_{33} = r - \delta - \frac{2rB^*}{K} = -\frac{\xi I}{B^*} - \frac{rB^*}{K} < 0.$$

It is obvious to see that $a_{11} < 0$. Therefore, we can conclude that

$$x = -(a_{11} + a_{22} + a_{33}) > 0.$$

Next, we show that $a_{11}a_{22} - a_{12}a_{21} > 0$. To begin, rewrite $a_{11} = -\mu N/S^*$ and $a_{21} = (\gamma + \mu)I^*/S^*$. Then, $a_{11}a_{22} - a_{12}a_{21}$

$$\begin{aligned} &= \frac{\mu N}{S^*} \left(\beta_2(B^*) \frac{S^* B^*}{\kappa + B^*} - S^* I^* \beta'_1(I^*) \right) + \frac{(\gamma + \mu)I^*}{S^*} (S^* \beta_1(I^*) + S^* I^* \beta'_1(I^*)) \\ &= \beta_2(B^*) \frac{S^* B^* \mu N}{S^* (\kappa + B^*)} + \frac{(\gamma + \mu)I^* S^* \beta_1(I^*)}{S^*} - \mu N I^* \beta'_1(I^*) + I^{*2} (\gamma + \mu) \beta'_1(I^*) \\ &= \beta_2(B^*) \frac{S^* B^* \mu N}{S^* (\kappa + B^*)} + \frac{(\gamma + \mu)I^* S^* \beta_1(I^*)}{S^*} - \beta'_1(I^*) I^* (\mu N - I^* (\gamma + \mu)) \\ &= \beta_2(B^*) \frac{S^* B^* \mu N}{S^* (\kappa + B^*)} + \frac{(\gamma + \mu)I^* S^* \beta_1(I^*)}{S^*} - \beta'_1(I^*) I^* \mu S^* > 0. \end{aligned}$$

The last inequality comes from subtracting Eq. (3.11) from Eq. (3.10). Now, using that $a_{11}a_{22} - a_{12}a_{21} > 0$, we can rewrite y as

$$y = A + \xi \frac{\beta'_2(B^*) S^* B^*}{(\kappa + B^*)}, \quad (\text{A.4})$$

where one can verify that A does not include the term $\beta'_2(B)$.

To ensure that $y > 0$, we introduce another assumption

$$\beta''_2(B) \leq \frac{2}{\kappa} \beta'_2(B). \quad (\text{A.5})$$

This condition provides an additional regulation for the indirect transmission function $\beta_2(B)$ that connects its first and second derivatives. It can be regarded as providing a lower bound for $\beta'_2(B)$ so that $\beta_2(B)$ remains biologically meaningful. In fact, Eq. (A.4) indicates that if $\beta'_2(B)$ approaches negative infinity, y would also be negative. Alternatively, condition (A.5) may be interpreted as setting an upper bound for $\beta''_2(B)$ such that it remains nonpositive (since $\beta'_2(B) \leq 0$), to represent possible saturation effects on the indirect transmission rate. Now, consider a Taylor series expansion of β_2 at $2B$,

$$\beta_2(2B) = \beta_2(B) + B\beta'_2(B) + \frac{B^2}{2}\beta''_2(\tilde{B}),$$

where $\tilde{B}$ is between B and $2B$. With condition (A.5), we obtain

$$\frac{B^2}{2}\beta_2''(\tilde{B}) \leq \frac{B^2}{\kappa}\beta_2'(\tilde{B}) \leq \frac{B^2}{\kappa}\beta_2'(B)$$

because $\beta_2(B)$ is a decreasing function and $B < \tilde{B}$. Based on this, we have

$$\kappa\beta_2(B) + B(\kappa + B)\beta_2'(B) > \kappa\beta_2(B) > 0,$$

which is equivalent to $a_{23} > 0$. It then follows that $y > 0$.

For $z = \xi(a_{11}a_{23} - a_{13}a_{21}) - a_{11}a_{22}a_{33} + a_{12}a_{21}a_{33}$, we know that $\xi a_{11}a_{23}$ is the only negative term in this equation, so we can show $z > 0$ by proving $\xi a_{11}a_{23} - a_{11}a_{22}a_{33} = a_{11}(\xi a_{23} - a_{33}a_{22}) > 0$. Since $a_{33} = r - \delta - \frac{2rB^*}{K} = -\frac{\xi I^*}{B^*} - \frac{rB^*}{K} < 0$, we can conclude that $a_{33}a_{22} - \xi a_{23} > 0$ by the following

$$\begin{aligned} & \left(-\frac{\xi I^*}{B^*} - \frac{rB^*}{K}\right) (S^*\beta_1' I^* + S\beta_1 - (\gamma + \mu)) - \xi \frac{\beta_2 S^* \kappa}{(\kappa + B^*)^2} \\ &= -\frac{\xi I^{*2}}{B} S\beta_1' - \frac{S^*\beta_1 \xi I^*}{B^*} + \frac{\xi(\gamma + \mu)I^*}{B^*} - \frac{rB^* S^* \beta_1' I^*}{K} - \frac{rB^* S^* \beta_1}{K} \\ & \quad + \frac{rB^*(\gamma + \mu)}{K} - \frac{\xi \beta_2 S^* \kappa}{(\kappa + B^*)^2} \\ &= -\frac{\xi I^{*2}}{B^*} S^* \beta_1' - \frac{S^*\beta_1 \xi I^*}{B^*} + \frac{\xi \beta_1 S^* I^*}{B^*} + \frac{\xi \beta_2 S^*}{(\kappa + B^*)} - \frac{rB^* S^* \beta_1' I^*}{K} - \frac{rB^* S^* \beta_1}{K} \\ & \quad + \frac{rB^* \beta_1 S^*}{K} + \frac{rB^{*2} \beta_2 S}{I^*(\kappa + B^*)K} - \frac{\xi \beta_2 S^* \kappa}{(\kappa + B^*)^2} \\ &> \left(\frac{\xi \beta_1 S^* I^*}{B^*} - \frac{\xi \beta_1 S^* I^*}{B^*}\right) + \left(\frac{\xi \beta_2 S^*}{\kappa + B^*} - \frac{\xi \beta_2 S^* \kappa}{(\kappa + B^*)^2}\right) \\ & \quad + \left(\frac{rB^* \beta_1 S^*}{K} - \frac{rB^* S^* \beta_1}{K}\right) \\ &> \frac{\xi \beta_2 S^* \kappa + \xi \beta_2 S^* B^*}{(\kappa + B^*)^2} - \frac{\xi \beta_2 S^* \kappa}{(\kappa + B^*)^2} > 0, \end{aligned}$$

which implies that $z > 0$. Lastly, we want to show $xy > z$. We observe that

$$\begin{aligned} xy - z &= (a_{11} + a_{22} + a_{33})[\xi a_{23} - a_{33}(a_{11} + a_{22})] \\ & \quad + \xi(a_{21}a_{13} - a_{11}a_{23}) + (a_{11} + a_{22})(a_{12}a_{21} - a_{11}a_{22}). \end{aligned} \quad (\text{A.6})$$

The last two terms are positive based on the former calculations. If the first term $a_{33}(a_{11} + a_{22}) - \xi a_{23} > 0$, then we can conclude that $xy - z > 0$. Note that we have already proved that $a_{33}a_{22} - \xi \frac{\beta_2 S^* \kappa}{(\kappa + B^*)^2} > 0$, hence

$$a_{33}(a_{11} + a_{22}) - \xi a_{23} \geq a_{33}a_{22} - \xi a_{23} = a_{33}a_{22} - \xi \frac{\beta_2 S^* \kappa}{(\kappa + B^*)^2} - \xi \frac{\beta_2' S^* B^*}{\kappa + B^*} > 0.$$

Therefore, it follows that $xy > z$. Consequently, the Routh–Hurwitz criteria yield that the endemic equilibrium is locally asymptotically stable.

Appendix B: Proof of Theorem 3.4

To show the global stability, we use the geometric approach by Li and Muldowney,³² which can be summarized in the following lemma.

Lemma B.1. *Let $\frac{dX}{dt} = f(X)$ be a dynamical system defined by a C^1 function $f : D \rightarrow \mathbf{R}^n$. Assume that $D \subset \mathbf{R}^n$ is simply connected and that there exists a compact absorbing set $K \subset D$. Additionally, assume that X^* is the only equilibrium of the dynamical system in D . Then, X^* is globally asymptotically stable in D provided that*

$$\bar{q}_2 = \limsup_{t \rightarrow \infty} \sup_{X_0 \in K} \frac{1}{t} \int_0^t m(Q(X(s, X_0))) ds < 0. \quad (\text{B.1})$$

Essentially, the condition $\bar{q}_2 < 0$ provides a Bendixson criterion in D . In the above equation, $m(A)$ is the Lozinskii measure of a matrix A with respect to a matrix norm; i.e.,

$$m(A) = \lim_{h \rightarrow 0^+} \frac{|I + hA| - 1}{h}$$

and Q is a function defined by $Q = P_f P^{-1} + P J^{[2]} P^{-1}$, where $P(X)$ is a matrix-valued C^1 function in D , P_f is the entry-wise derivative of P along the direction of f , and $J^{[2]}$ is the second additive compound matrix associated with the Jacobian matrix of the system. Particularly, for a 3×3 matrix $A = [a_{ij}]$, the second additive compound matrix is defined as

$$A^{[2]} = \begin{pmatrix} a_{11} + a_{22} & a_{23} & -a_{13} \\ a_{32} & a_{11} + a_{33} & a_{12} \\ -a_{31} & a_{21} & a_{22} + a_{33} \end{pmatrix}.$$

We refer to Ref. 33 for a survey of general compound matrices.

We can now prove Theorem 3.4. First, Theorem 3.1 states that the DFE is unstable when $R_0 > 1$. Since the DFE is located on the boundary of the domain Ω , it follows that the system is uniformly persistent when $R_0 > 1$. The compactness of Ω and the uniform persistence of the system implies that there exists a compact absorbing set. Meanwhile, Theorem 3.2 ensures that there is a unique endemic equilibrium in the interior of Ω when $R_0 > 1$.

Based on Lemma B.1, we only need to verify the condition (B.1) for our system. The Jacobian of the system without the equations for R (since $R = N - S - I$) is given by

$$\begin{bmatrix} -\beta_1(I)I - \frac{\beta_2(B)B}{\kappa + B} - \mu & -S\beta_1(I) - SI\beta'_1(I) & -\frac{\beta_2(B)S\kappa}{(\kappa + B)^2} - \frac{\beta'_2(B)SB}{(\kappa + B)} \\ \beta_1(I)I + \frac{\beta_2(B)B}{\kappa + B} & S\beta_1(I) + SI\beta'_1(I) - (\gamma + \mu) & \frac{\beta_2(B)S\kappa}{(\kappa + B)^2} + \frac{\beta'_2(B)SB}{(\kappa + B)} \\ 0 & \xi & r - \delta - \frac{2rB}{K} \end{bmatrix}.$$

For ease of notation, we will denote $a = S\beta_1(I) + SI\beta'_1(I)$, $b = \frac{\beta_2(B)S\kappa}{(\kappa+B)^2} + \frac{\beta'_2(B)SB}{(\kappa+B)}$, and $c = \beta_1(I)I + \frac{\beta_2(B)B}{\kappa+B}$. The second additive compound matrix is

$$J^{[2]} = \begin{bmatrix} a - c - 2\mu - \gamma & b & b \\ \xi & -c - \mu + r - \delta - \frac{2rB}{K} & -a \\ 0 & c & a - \gamma - \mu + r - \delta - \frac{2rB}{K} \end{bmatrix}.$$

Now we take $P = \text{diag}[1, \frac{I}{B}, \frac{I}{B}]$ so that $P_f P^{-1} = \text{diag}[0, \frac{I'}{I} - \frac{B'}{B}, \frac{I'}{I} - \frac{B'}{B}]$. Then, we can write $Q = P_f P^{-1} + P J^{[2]} P^{-1}$ as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

where

$$Q_{11} = a - c - 2\mu - \gamma,$$

$$Q_{12} = \begin{bmatrix} \frac{B}{I}b, & \frac{B}{I}b \end{bmatrix},$$

$$Q_{21} = \begin{bmatrix} \frac{\xi I}{B}, & 0 \end{bmatrix}^T,$$

$$Q_{22} = \begin{bmatrix} -c - \mu + r - \delta - \frac{2rB}{K} - \frac{B'}{B} + \frac{I'}{I} & -a \\ c & a - \gamma - \mu + r - \delta - \frac{2rB}{K} - \frac{B'}{B} + \frac{I'}{I} \end{bmatrix}.$$

Define a vector norm for any $(x, y, z) \in \mathbf{R}^3$ as $|(x, y, z)| = \max(|x|, |y| + |z|)$. The Lozinskii measure induced by this norm is $m(Q) = \sup(g_1, g_2)$ where

$$g_1 = m_1(Q_{11}) + |Q_{12}|, \quad g_2 = m_1(Q_{22}) + |Q_{21}|$$

and where in this case $|Q|$ is the L_1 matrix norm, and m_1 is the Lozinskii measure with respect to the L_1 norm. In particular, for our system, we obtain

$$g_1 = \frac{B}{I}b + a - \gamma - 2\mu - \beta_1 I - \frac{\beta_2 B}{\kappa + B},$$

$$g_2 = \frac{\xi I}{B} + r - \delta - \frac{2rB}{K} - \frac{B'}{B} + \frac{I'}{I} - \mu + \max(0, a + |a| - \gamma).$$

Note that $\frac{I'}{I} = \beta_1(I)S + \frac{\beta_2(B)SB}{I(\kappa+B)} - \gamma - \mu$. Also, because $\frac{\kappa}{\kappa+B} \leq 1$, we know that $\frac{\beta_2(B)SB\kappa}{I(\kappa+B)^2} \leq \frac{\beta_2(B)SB}{I(\kappa+B)}$. Therefore,

$$\begin{aligned} g_1 &= \frac{B\beta_2(B)S\kappa}{I(\kappa+B)^2} + \beta_1(I)S - \gamma - 2\mu + \frac{\beta'_2(B)SB^2}{I(\kappa+B)} - \beta_1(I)I + \beta'_1(I)IS - \frac{\beta_2(B)B}{\kappa+B} \\ &\leq \frac{I'}{I} - \mu + \frac{\beta'_2(B)SB^2}{I(\kappa+B)} - \beta_1(I)I + \beta'_1(I)IS - \frac{\beta_2(B)B}{\kappa+B} \\ &\leq \frac{I'}{I} - \mu. \end{aligned} \tag{B.2}$$

Then, under the assumption that $N \leq \frac{\gamma}{2\beta_1(0)}$, and substituting $\frac{B'}{B} = r - \delta + \frac{\xi I}{B} - \frac{2rB}{K}$, we obtain

$$g_2 = \frac{I'}{I} - \mu - \frac{rB}{K} + \max(0, a + |a| - \gamma) \leq \frac{I'}{I} - \mu, \tag{B.3}$$

because if $a > 0$, then $2a - \gamma = 2S(\beta_1(I) + I\beta'_1(I)) - \gamma \leq 2N\beta_1(0) - \gamma \leq 0$. This uses the fact that N is an upper bound on S , 0 is a lower bound on B , and $\beta_1(0)$ is an upper bound for $\beta_1(I)$ since it is a decreasing function. Meanwhile, if $a \leq 0$, then the inequality in (B.3) automatically holds.

Therefore, based on (B.2) and (B.3), it is clear that $m(Q) \leq \frac{I'}{I} - \mu$. Since $0 \leq I(t) \leq N$, for sufficiently large t , $\frac{\ln(I(t)) - \ln(I(0))}{t} \leq \frac{\mu}{2}$ will be true. Therefore,

$$\frac{1}{t} \int_0^t m(Q) ds \leq \frac{1}{t} \int_0^t \left(\frac{I'}{I} - \mu \right) ds = \frac{\ln(I(t)) - \ln(I(0))}{t} - \mu \leq -\frac{\mu}{2},$$

which implies that

$$\bar{q}_2 = \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t m(Q) ds \leq -\frac{\mu}{2} < 0.$$

Thus, by Lemma B.1, the endemic equilibrium is globally asymptotically stable.