Tenth MSU Conference on Differential Equations and Computational Simulations.
Electronic Journal of Differential Equations, Conference 23 (2016), pp. 87–117.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

OPTIMAL CONTROL IN MULTI-GROUP COUPLED WITHIN-HOST AND BETWEEN-HOST MODELS

ERIC NUMFOR, SOUVIK BHATTACHARYA,
MAIA MARTCHEVA, SUZANNE LENHART

Abstract. We formulate and then analyze a multi-group coupled within-host model of ODEs and between-host model of ODE and first-order PDEs, using the Human Immunodeficiency Virus (HIV) for illustration. The basic reproduction number of the multi-group coupled epidemiological model is derived, steady states solutions are calculated and stability analysis of equilbria is investigated. An optimal control problem for our model with drug treatment on the multi-group within-host system is formulated and analyzed. Ekeland's principle is used in proving existence and uniqueness of an optimal control pair. Numerical simulations based on the semi-implicit finite difference schemes and the forward-backward sweep iterative method are obtained.

1. Introduction

The two key features in infectious diseases are the transmission between hosts and the immunological process at the individual host level. Understanding how the two features influence each other can be assisted through modeling. Linking components of the immune system with the compartments of the epidemic model leads to a two-scale model. Much of the work on such "linked" models deal with the two levels separately, making "decoupling" assumptions [1].

Despite advancements in the study of epidemiological, within-host and immunological models, the outbreak of some diseases cannot still be predicted. This dilemma may be attributed to the fact that most modeling approaches are either restricted to epidemiological or immunological formulations [23]. Current research focuses on the comprehensive modeling approach, called immuno-epidemiological modeling, which investigates the influence of population immunity on epidemiological patterns, translates individual characteristics such as immune status and pathogen load to population level and traces their epidemiological significance [12, 24, 30]. Several immuno-epidemiological models have been used to study the relationship between transmission and virulence [5, 14, 15, 19, 20, 21]. Some of these models deal with the two processes separately by making decoupling assumptions. Gilchrist and Sasaki [20] used the nested approach to model host-parasite

2010 *Mathematics Subject Classification.* 34D20, 35L02, 49K15, 49K15.
Key words and phrases. Multi-group; within-host dynamics; between-host dynamics; equilibria; basic reproduction number; stability; optimal control.

Published March 21, 2016.

coevolution in which the within-host model is independent of the between-host but the parameters of the between-host model are expressed in terms of dependent variables of the within-host model. Also, Feng et al. [14] investigated a coupled within-host and between-host model of Toxoplasma gondii linked via the environment. Numfor et al. [34] set a framework for optimal control of coupled within-host and between-host models.

Our goals are to use a multi-group within-host model coupled with a epidemiology model to capture the impact on the epidemic of giving treatment to individuals, and investigate mathematically such a coupled ODE/PDE system (well-posedness and optimal control). Our general approach in immuno-epidemiological modeling involves a nesting approach [21, 34] whereby the within-host model is nested within the epidemiological model by linking the dynamics of the within-host model to the additional host mortality, recovery and transmission rates of the infection. The within-host and between-host models are also linked via a structural variable.

This work will have the first results on formulating this multi-group two-scale model in a careful mathematical framework and the first results on optimal control of such a multi-group model. We emphasize the novelty of mathematical results, as well as the importance of the epidemiological and immunological results. To curtail the proliferation of free virus at the within-host level, we introduce two functions, representing transmission and virion production suppressing drugs. Our goal is to use optimal control techniques in the coupled model to minimize free virus at the within-host level and infectious individuals at the population level, while minimizing the cost of implementing the controls (this may include toxicity effects). Optimal control of first-order partial differential equations is done differently than optimal control of parabolic PDEs due to the lack of regularity of solutions to the first-order PDEs. The steps in justifying the optimal control results are quite different and we use Ekeland's principle [13] to get the existence of an optimal control.

The remainder of the work in this paper is organized as follows. In section 2, we present our multi-group within-host and between-host models. The multi-group within-host model is independent of the between-host model, but the between-host model is linked to the within-host via coefficients and a structural variable. In section 3, we establish the boundedness of state solutions to the within-host model, and existence and uniqueness of solutions to the between-host model is established. An explicit expression for the basic reproduction number of the epidemiological model is derived, steady state solutions are calculated and stability analysis of equilibrium points is studied. In section 4, an optimal control problem with drug treatment on the within-host model is formulated. Lipschitz properties of state and adjoint solutions in terms of the control functions are shown. The differentiability of the solution map and existence of adjoints solutions are established. The lower semi-continuity of the objective functional with respect to L^1 convergence is established. The last part of section 4 is devoted to the existence of a unique optimal control pair, which is obtained with the use of Ekeland's principle. Numerical simulations based on the semi-implicit finite difference schemes and a forward-backward sweep iterative method [2, 28] will be studied in section 5, and our concluding remarks presented in section 5.

2. Multi-group Within-host and Between-host Models

In our multi-group within-host and between-host model, we assume that all individuals in the population exhibit different immunological dynamics upon infection. Since individuals with stronger immune systems respond better to treatment in the case of antiretroviral therapy for the human immunodeficiency virus (HIV), and the optimum viral load required for shedding depends on the strength of the cytotoxic T lymphocyte (CTL) response of the particular host, we focus only on two classes of individuals with different immunological characteristics and viral load. Thus, the within-host dynamics of pathogen for each individual of group j is

$$\frac{dx_j}{d\tau} = r - \beta_j v_j(\tau)x_j(\tau) - \mu x_j(\tau), \quad x_j(0) = x_j^0 \tag{2.1}$$

$$\frac{dy_j}{d\tau} = \beta_j v_j(\tau)x_j(\tau) - d_j y_j(\tau), \quad y_j(0) = y_j^0 \tag{2.2}$$

$$\frac{dv_j}{d\tau} = \gamma_j d_j y_j(\tau) - (\delta_j + s_j)v_j(\tau) - \hat{\beta}_j v_j(\tau)x_j(\tau), \quad v_j(0) = v_j^0, \tag{2.3}$$

where $j = 1, 2$ defines the two classes of individuals with different immunological characteristics and viral load. In the model, x_j defines the number of healthy cells in the jth immunological class which is being produced at a constant rate r and die at rate μ. The growth and death rates of healthy cells are assumed to be the same for all individuals in all immunological classes. These healthy cells come in contact with free virus v_j at rate β_j and become infected cells y_j, with $\hat{\beta}_j$ being the binding rate of the virus to healthy cells. The infected cells in the jth group die at rate d_j and each produce γ_j virions at bursting. The clearance and shedding rates of the virus are δ_j and s_j, respectively.

The epidemiological model is divided into two classes; individuals in each epidemiological class exhibits different immunological characteristics. We denote the number of susceptible individuals at time t by $S(t)$, and the density of infected individual structured by chronological time t and age-since-infection τ by $i_j(\tau, t)$, where $j = 1, 2$. Individuals in each group exhibit the same immunological characteristics, but individuals in different groups exhibit different immunological characteristics and viral load. Our multi-group epidemiological (or between-host) model is:

$$\frac{dS}{dt} = \Lambda - \frac{S}{N}\sum_{j=1}^{2}\int_0^A c_j s_j v_j(\tau)i_j(\tau, t)d\tau - m_0 S \quad in \quad (0, T) \tag{2.4}$$

$$\frac{\partial i_1}{\partial t} + \frac{\partial i_1}{\partial \tau} = -m(v_1(\tau))i_1(\tau, t) \quad in \quad (0, A) \times (0, T) \tag{2.5}$$

$$i_1(0, t) = p_1\frac{S}{N}\int_0^A c_1 s_1 v_1(\tau)i_1(\tau, t)d\tau + p_1\frac{S}{N}\int_0^A c_2 s_2 v_2(\tau)i_2(\tau, t)d\tau \tag{2.6}$$

$$\frac{\partial i_2}{\partial t} + \frac{\partial i_2}{\partial \tau} = -m(v_2(\tau))i_2(\tau, t) \quad in \quad (0, A) \times (0, T) \tag{2.7}$$

$$i_2(0, t) = p_2\frac{S}{N}\int_0^A c_1 s_1 v_1(\tau)i_1(\tau, t)d\tau + p_2\frac{S}{N}\int_0^A c_2 s_2 v_2(\tau)i_2(\tau, t)d\tau \tag{2.8}$$

$$i_1(\tau, 0) = i_1^0(\tau), \quad i_2(\tau, 0) = i_2^0(\tau). \tag{2.9}$$

In the epidemiological model, $m(v_j(\tau))$ is the death rate of infected hosts (a function of viral load) in the jth class, Λ is the recruitment rate of susceptible individuals, $m_0 = m(0)$ is the death rate of susceptible individuals and p_j is the probability that

an individual who is infected has immunological behavior similar to individuals in the jth class, with $p_1+p_2=1$. The transmission rate is assumed to be proportional to the viral load of infected individuals in the jth group, calculated by integrating with respect to τ, $\int_0^A (c_1 s_1 v_1(\tau) i_1(\tau,t) + c_2 s_2 v_2(\tau) i_2(\tau,t))d\tau$, where c_j is the contact rate between susceptible and infected individuals. Thus, the new infections of the population in group j at time t, denoted by $i_j(0,t)$, depends on the age distribution of the population at time t, as determined by the integral of $i_j(\tau,t)$ over all ages, weighted with the specific transmission rate $\tilde{\beta}_j(\tau) = c_j s_j v_j(\tau)$. The number of susceptible and infectious individuals in the population at time $t=0$ are given by $S(0)=S_0>0$ and $i_j(\tau,0)=i_j^0(\tau)$, respectively. Thus, $i_j(\tau,0)$ is the initial age distribution of infectious individuals in group j, with i_j^0 being a known nonnegative function of age-since-infection, τ. The total population of infectious individuals of each group from birth to maximal age-since-infection, A, is defined as

$$I_1(t)=\int_0^A i_1(\tau,t)d\tau \quad \text{and} \quad I_2(t)=\int_0^A i_2(\tau,t)d\tau,$$

and the total population size of individuals in the population at time t is $N(t)=S(t)+I_1(t)+I_2(t)$. For the sake of introduction to our method, we assume the simplest form for the mortality function [11], $m(v_j)$, as

$$m(v_j(\tau)) = m_0 + \mu_j v_j(\tau),$$

so that in the absence of the virus, individuals die naturally at rate m_0. The term $\mu_j v_j(\tau)$ gives the additional host mortality in group j due to the virus.

3. Existence of solution, equilibria and stability analysis of the epidemiological model

3.1. **Existence of solution.** We use a result from [34] which applies the fixed point argument to obtain an existence and uniqueness of solution to our coupled model. To do this, we use the method of integrating factors on the differential equation (2.4), and integrating the differential equations (2.5) and (2.7) along the characteristic line $\tau - t =$ constant and considering cases where $\tau > t$ and $\tau < t$ to obtain a representation formula for the solution to the epidemiological model.

In using the fixed point argument for the existence of solution, we define our state solution space as

$$\begin{aligned} X = \Big\{ &(S,i_1,i_2) \in L^\infty(0,T)\times (L^\infty(0,T;L^1(0,A)))^2 | S(t)\geq \varepsilon>0,\ i_1(\tau,t)\geq 0,\\ &i_2(\tau,t)\geq 0,\ \sup_t S(t)<\infty,\ \sup_t \int_0^A i_1(\tau,t)d\tau<\infty,\\ &\text{and } \sup_t \int_0^A i_2(\tau,t)d\tau<\infty \ a.e.\ t\Big\}, \end{aligned}$$

where $\varepsilon=\min\{S_0, \frac{\Lambda}{m_0+\tilde{\alpha}}\}$ and $\tilde{\alpha}$ is a positive number that satisfies the inequality $\tilde{\alpha}\geq C(c_1s_1+c_2s_2)>0$. The constant C is a bound for v_j. Now, we define a map

$$\mathcal{L}: X\to X, \quad \mathcal{L}(S,i_1,i_2)=(L_1(S,i_1,i_2),L_2(S,i_1,i_2),L_3(S,i_1,i_2)),$$

where

$$\begin{aligned} &L_1(S,i)(t)\\ &= S_0 e^{(m_0+\tilde{\alpha})t} + \frac{\Lambda}{m_0+\tilde{\alpha}}(1-e^{-(m_0+\tilde{\alpha})t})\\ &\quad + \int_0^t e^{-(m_0+\tilde{\alpha})(t-s)} S(s)\Big(\tilde{\alpha} - \frac{1}{N(s)}\sum_{j=1}^{2}\int_0^A c_j s_j v_j(\tau) i_j(\tau,s)d\tau\Big)ds \end{aligned} \tag{3.1}$$

$$\begin{aligned} &L_2(S,i)(\tau,t)\\ &= \begin{cases} p_1 \frac{S(t-\tau)}{N(t-\tau)} e^{-\int_0^\tau m(v_1(s))ds} \sum_{j=1}^{2}\int_0^A c_j s_j v_j(s) i_j(s,t-\tau)ds, & \tau < t\\ i_1^0(\tau-t) e^{-\int_0^t m(v_1(\tau-t+s))ds}, & \tau > t \end{cases} \end{aligned} \tag{3.2}$$

$$\begin{aligned} &L_3(S,i)(\tau,t)\\ &= \begin{cases} p_2 \frac{S(t-\tau)}{N(t-\tau)} e^{-\int_0^\tau m(v_2(s))ds} \sum_{j=1}^{2}\int_0^A c_j s_j v_j(s) i_j(s,t-\tau)ds, & \tau < t\\ i_2^0(\tau-t) e^{-\int_0^t m(v_2(\tau-t+s))ds}, & \tau > t. \end{cases} \end{aligned} \tag{3.3}$$

where $L_1(S,i)(t)$ is a representation formula for the solution to the differential equation

$$\frac{dS}{dt} + \tilde{\alpha}S(t) = \Lambda + \tilde{\alpha}S(t) - \frac{S(t)}{N(t)}\sum_{j=1}^{2}\int_0^A c_j s_j v_j(\tau) i_j(\tau,t)d\tau - m_0 S(t).$$

This differential equation is equivalent to (2.4).

The following assumptions will be useful in establishing a Lipschitz property for the within-host and between-host state solutions in terms of control functions:

(1) S_0, m_0, Λ, c_j and s_j are positive constants,
(2) $m(s)$ is non-negative and Lipschitz continuous,
(3) $i_j^0(\tau)$ is non-negative for all $\tau \in (0,A)$,
(4) $\int_0^A i_j^0(\tau)d\tau \le M$ and $0 < S_0 \le M$.

Remark 3.1. Starting with positive initial data, state solutions of the multi-group model stay positive for all $\tau > 0$, and are bounded in finite time [34].

Theorem 3.2. *For $T < \infty$, there exists a unique nonnegative solution (S, i_1, i_2) to the epidemiological system* (2.4)–(2.9).

Proof. We show that the map $\mathcal{L}$ maps X into itself, and that $\mathcal{L}$ admits a unique fixed point by defining an iterative sequence [25, 32]. For details, see Numfor et al. [34]. □

3.2. **Basic reproduction number and equilibria.** We derive the basic reproduction number for our multi-group coupled epidemiological model, and investigate the existence of equilibria. In deriving the basic reproduction number, $\mathcal{R}_0$, we compute the disease-free equilibrium, linearize the system around the disease-free equilibrium and determine conditions for its stability. Now, we consider solutions near the disease-free equilibrium $(S^*, i_1^*(\tau), i_2^*(\tau)) = (\frac{\Lambda}{m_0}, 0, 0)$ by setting

$$x(t) = S(t) - S^*, \quad y_1(\tau,t) = i_1(\tau,t), \quad y_2(\tau,t) = i_2(\tau,t).$$

Substituting the perturbed solutions into equations (2.4)–(2.9), we obtain the following linearized system:

$$\frac{dx}{dt} = -\sum_{j=1}^{2}\int_0^A c_j s_j v_j(\tau) y_j(\tau,t)d\tau - m_0 x(t) \tag{3.4}$$

$$\frac{\partial y_1}{\partial t} + \frac{\partial y_1}{\partial \tau} = -m(v_1(\tau))y_1(\tau,t) \tag{3.5}$$

$$y_1(0,t) = p_1\Big(\int_0^A c_1 s_1 v_1(\tau) y_1(\tau,t)d\tau + \int_0^A c_2 s_2 v_2(\tau) y_2(\tau,t)d\tau\Big) \tag{3.6}$$

$$\frac{\partial y_2}{\partial t} + \frac{\partial y_2}{\partial \tau} = -m(v_2(\tau))y_2(\tau,t) \tag{3.7}$$

$$y_2(0,t) = p_2\Big(\int_0^A c_1 s_1 v_1(\tau) y_1(\tau,t)d\tau + \int_0^A c_2 s_2 v_2(\tau) y_2(\tau,t)d\tau\Big). \tag{3.8}$$

We seek solutions to the first-order partial differential equations (3.5) and (3.7) of the form

$$y_1(\tau,t) = \bar{y}_1(\tau)e^{\lambda t} \quad \text{and} \quad y_2(\tau,t) = \bar{y}_2(\tau)e^{\lambda t},$$

where λ is either a real or complex number. Substituting these solutions into equations (3.5)–(3.8), we have the following eigenfunction problem

$$\frac{d\bar{y}_1(\tau)}{d\tau} = -(\lambda + m(v_1(\tau)))\bar{y}_1(\tau) \tag{3.9}$$

$$\bar{y}_1(0) = p_1\Big(\int_0^A c_1 s_1 v_1(\tau)\bar{y}_1(\tau)d\tau + \int_0^A c_2 s_2 v_2(\tau)\bar{y}_2(\tau)d\tau\Big) \tag{3.10}$$

$$\frac{d\bar{y}_2(\tau)}{d\tau} = -(\lambda + m(v_2(\tau)))\bar{y}_2(\tau) \tag{3.11}$$

$$\bar{y}_2(0) = p_2\Big(\int_0^A c_1 s_1 v_1(\tau)\bar{y}_1(\tau)d\tau + \int_0^A c_2 s_2 v_2(\tau)\bar{y}_2(\tau)d\tau\Big). \tag{3.12}$$

The solutions to equations (3.9) and (3.11) are

$$\bar{y}_1(\tau) = \bar{y}_1(0)e^{-\lambda\tau}e^{-\int_0^\tau m(v_1(s))ds}, \quad \bar{y}_2(\tau) = \bar{y}_2(0)e^{-\lambda\tau}e^{-\int_0^\tau m(v_2(s))ds},$$

so that the initial conditions (3.10) and (3.12) become

$$\bar{y}_1(0) = p_1\sum_{j=1}^{2} c_j s_j \bar{y}_j(0)\int_0^A v_j(\tau)e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau$$

$$\bar{y}_2(0) = p_2\sum_{j=1}^{2} c_j s_j \bar{y}_j(0)\int_0^A v_j(\tau)e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau.$$

The eigenfunction problem (3.9)–(3.12) has a non-trivial solution if, and only if,

$$(p_1J_1 - 1)(p_2J_2 - 1) - p_1p_2J_1J_1 = 0,$$

where $J_\ell = c_\ell s_\ell \int_0^A v_\ell(\tau)e^{-\lambda\tau}e^{-\int_0^\tau m(v_\ell(s))ds}d\tau$. This gives

$$1 = p_1J_1 + p_2J_2 \equiv \sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau. \tag{3.13}$$

The right-hand side of equation (3.13) is a function of λ, which we denote by $G(\lambda)$, where

$$G(\lambda) = \sum_{j=1}^{2} \int_0^A p_j c_j s_j v_j(\tau) e^{-\lambda\tau} e^{-\int_0^\tau m(v_j(s))ds} d\tau, \tag{3.14}$$

so that $G(\lambda) = 1$ is a characteristic equation that will be used to study stability of the disease-free equilibrium. We define the basic reproduction number, $\mathcal{R}_0$, of the epidemiological (or linked) model as $\mathcal{R}_0 = G(0)$ [9, 29, 31, 36, 37] so that

$$\mathcal{R}_0 = \sum_{j=1}^{2} \int_0^A p_j c_j s_j v_j(\tau) e^{-\int_0^\tau m(v_j(s))ds} d\tau, \tag{3.15}$$

where $\pi_j(\tau) = e^{-\int_0^\tau m(v_j(s))ds}$ is the probability of survival in the infected class of group j from onset of infection to age-since-infection, τ, and p_j is the probability that an individual who is infected has immunological behavior similar to individuals in the jth class.

Theorem 3.3. *The epidemiological model has a unique endemic equilibrium, $(S^*, i_1^*(\tau), i_2^*(\tau))$, if $\mathcal{R}_0 > 1$.*

Proof. We set the time derivatives of the epidemiological model to zero. This gives:

$$0 = \Lambda - \frac{S}{N} \sum_{j=1}^{2} \int_0^A c_j s_j v_j(\tau) i_j(\tau) d\tau - m_0 S \tag{3.16}$$

$$\frac{di_j(\tau)}{d\tau} = -m(v_j(\tau)) i_j(\tau) \tag{3.17}$$

$$i_j(0) = p_j \frac{S}{N} \sum_{k=1}^{2} \int_0^A c_k s_k v_k(\tau) i_k(\tau) d\tau. \tag{3.18}$$

To derive the endemic equilibrium, we solve the differential equation (3.17) to have

$$i_j^*(\tau) = i_j^*(0) e^{-\int_0^\tau m(v_j(s))ds}. \tag{3.19}$$

Next, substituting the expression for $i_j^*(\tau)$ in (3.16) yields

$$0 = \Lambda - \frac{S^*}{N^*} \sum_{j=1}^{2} \int_0^A c_j s_j v_j(\tau) i_j^*(0) e^{-\int_0^\tau m(v_j(s))ds} d\tau - m_0 S^*. \tag{3.20}$$

From (3.18), (3.19) and (3.20), we obtain $i_j^*(0)$ as

$$i_j^*(0) = p_j(\Lambda - m_0 S^*).$$

Since the total population at equilibrium is $N^* = S^* + \int_0^A i_1^*(\tau)d\tau + \int_0^A i_2^*(\tau)d\tau$, we obtain $N^* = \Lambda\xi + (1 - m_0\xi)S^*$, where

$$\xi = p_1 \int_0^A e^{-\int_0^\tau m(v_1(s))ds} d\tau + p_2 \int_0^A e^{-\int_0^\tau m(v_2(s))ds} d\tau.$$

Now, from (3.16), we have

$$\frac{S^*}{N^*} = \frac{i_j^*(0)}{p_j(\Lambda - m_0 S^*)\mathcal{R}_0} = \frac{1}{\mathcal{R}_0},$$

so that

$$S^* = \frac{\Lambda\xi}{\mathcal{R}_0 - 1 + m_0\xi} \quad \text{and} \quad i_j^*(\tau) = \frac{p_j\Lambda(\mathcal{R}_0 - 1)e^{-\int_0^\tau m(v_j(s))ds}}{\mathcal{R}_0 - 1 + m_0\xi}.$$

Hence, the endemic equilibrium is $(S^*, i_1^*(\tau), i_2^*(\tau))$, where

$$\begin{aligned}&(S^*, i_1^*(\tau), i_2^*(\tau))\\ &= \Big(\frac{\Lambda\xi}{\mathcal{R}_0 - 1 + m_0\xi}, \frac{p_1\Lambda(\mathcal{R}_0 - 1)e^{-\int_0^\tau m(v_1(s))ds}}{\mathcal{R}_0 - 1 + m_0\xi}, \frac{p_2\Lambda(\mathcal{R}_0 - 1)e^{-\int_0^\tau m(v_2(s))ds}}{\mathcal{R}_0 - 1 + m_0\xi}\Big),\end{aligned}$$

which exists if $\mathcal{R}_0 > 1$. □

3.3. **Stability analysis.** To study the local stability of equilibria, we linearize the model around each of the equilibrium points, and consider an exponential solution to the linearized system.

Theorem 3.4. *The disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 < 1$ and unstable if $\mathcal{R}_0 > 1$.*

Proof. If $\lambda \in \mathbb{R}$, then from equation (3.14), $G'(\lambda) < 0$, since v_j is non-negative and bounded. Thus, G is a decreasing function of λ. Therefore, there exists a unique positive solution to the characteristic equation $G(\lambda) = 1$ when $\mathcal{R}_0 = G(0) > 1$, since $G(\lambda) \to 0$ as $\lambda \to \infty$. Hence, the disease-free equilibrium is unstable when $\mathcal{R}_0 > 1$.

When $\mathcal{R}_0 = G(0) < 1$, there exists a unique negative solution to the characteristic equation $G(\lambda) = 1$, since $G(\lambda) \to +\infty$ as $\lambda \to -\infty$. Next, we assume that λ is complex and let $\lambda = \eta_1 + i\eta_2$ be an arbitrary complex solution (if it exists) to the characteristic equation $G(\lambda) = 1$. Then

$$\begin{aligned}1 &= |G(\eta_1 + i\eta_2)|\\ &\le \sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)e^{-\eta_1\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau =: G(\Re(\lambda)).\end{aligned}$$

If $\Re(\lambda) \ge 0$, then $1 \le G(\Re(\lambda)) \le G(0) = \mathcal{R}_0 < 1$, which is absurd. Thus, all roots of $G(\lambda) = 1$ have negative real parts when $\mathcal{R}_0 < 1$. Hence the disease-free equilibrium is locally asymptotically stable when $\mathcal{R}_0 < 1$. □

Theorem 3.5. *The disease-free equilibrium is globally stable if $\mathcal{R}_0 < 1$.*

The proof of the above theorem follows as in Numfor et al. [34, Theorem 2.5].

Theorem 3.6. *The endemic equilibrium*

$$\begin{aligned}&(S^*, i_1^*(\tau), i_2^*(\tau))\\ &= \Big(\frac{\Lambda\xi}{\mathcal{R}_0 - 1 + m_0\xi}, \frac{p_1\Lambda(\mathcal{R}_0 - 1)e^{-\int_0^\tau m(v_1(s))ds}}{\mathcal{R}_0 - 1 + m_0\xi}, \frac{p_2\Lambda(\mathcal{R}_0 - 1)e^{-\int_0^\tau m(v_2(s))ds}}{\mathcal{R}_0 - 1 + m_0\xi}\Big)\end{aligned}$$

is locally asymptotically stable if $\mathcal{R}_0 > 1$ and the maximal age of infection, A, is sufficiently large.

Proof. We consider solutions near the endemic equilibrium by setting

$$x(t) = S(t) - S^*, \quad y_1(\tau, t) = i_1(\tau, t) - i_1^*(\tau), \quad y_2(\tau, t) = i_2(\tau, t) - i_2^*(\tau),$$

so that the total population is $N(t) = N^* + n(t)$, where

$$n(t) = x(t) + \int_0^A y_1(\tau,t)d\tau + \int_0^A y_2(\tau,t)d\tau, \quad N^* = S^* + \int_0^A i_1^*(\tau)d\tau + \int_0^A i_2^*(\tau)d\tau.$$

Substituting the perturbed solutions into (2.4)–(2.9), we have the linearized system

$$\begin{aligned}
\frac{dx}{dt} &= -\frac{x(t)}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau + \frac{S^*}{N^*}\frac{n(t)}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad - \frac{x(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau + \frac{S^*}{N^*}\frac{n(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau \\
&\quad - \frac{S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau) y_1(\tau,t)d\tau - \frac{S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau) y_2(\tau,t)d\tau - m_0 x
\end{aligned} \tag{3.21}$$

$$\frac{\partial y_1}{\partial t} + \frac{\partial y_1}{\partial \tau} = -m(v_1(\tau))y_1(\tau,t) \tag{3.22}$$

$$\begin{aligned}
y_1(0,t) &= \frac{p_1 x}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau - \frac{p_1 S^*}{N^*}\frac{n}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad + \frac{p_1 S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau) y_1(\tau,t)d\tau + \frac{p_1 S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau) y_2(\tau,t)d\tau \\
&\quad + \frac{p_1 x(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau - \frac{p_1 S^*}{N^*}\frac{n(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau
\end{aligned} \tag{3.23}$$

$$\frac{\partial y_2}{\partial t} + \frac{\partial y_2}{\partial \tau} = -m(v_2(\tau))y_2(\tau,t) \tag{3.24}$$

$$\begin{aligned}
y_2(0,t) &= \frac{p_2 x}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau - \frac{p_2 S^*}{N^*}\frac{n}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad + \frac{p_2 S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau) y_1(\tau,t)d\tau + \frac{p_2 S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau) y_2(\tau,t)d\tau \\
&\quad + \frac{p_2 x(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau - \frac{p_2 S^*}{N^*}\frac{n(t)}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau.
\end{aligned} \tag{3.25}$$

Next, we seek solutions to (3.21)–(3.25) of the form

$$x(t) = \bar{x}e^{\lambda t}, \quad y_1(\tau,t) = \bar{y}_1(\tau)e^{\lambda t}, \quad y_2(\tau,t) = \bar{y}_2(\tau)e^{\lambda t}.$$

This gives

$$\begin{aligned}
\lambda\bar{x} &= -\frac{\bar{x}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau + \frac{S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad - \frac{\bar{x}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau + \frac{S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau \\
&\quad - \frac{S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau) \bar{y}_1(\tau)d\tau - \frac{S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau) \bar{y}_2(\tau)d\tau - m_0\bar{x}
\end{aligned} \tag{3.26}$$

$$\frac{d\bar{y}_1(\tau)}{d\tau} = -(\lambda + m(v_1(\tau)))\bar{y}_1(\tau) \tag{3.27}$$

$$\begin{aligned}
\bar{y}_1(0) &= \frac{p_1\bar{x}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau - \frac{p_1 S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad + \frac{p_1\bar{x}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau - \frac{p_1 S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau \\
&\quad + \frac{p_1 S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau)\bar{y}_1(\tau)d\tau + \frac{p_1 S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau)\bar{y}_2(\tau)d\tau
\end{aligned} \tag{3.28}$$

$$\frac{d\bar{y}_2(\tau)}{d\tau} = -(\lambda + m(v_2(\tau)))\bar{y}_2(\tau) \tag{3.29}$$

$$\begin{aligned}
\bar{y}_2(0) &= \frac{p_2\bar{x}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau - \frac{p_2 S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_1 s_1 v_1(\tau) i_1^*(\tau)d\tau \\
&\quad + \frac{p_2\bar{x}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau - \frac{p_2 S^*}{N^*}\frac{\bar{n}}{N^*}\int_0^A c_2 s_2 v_2(\tau) i_2^*(\tau)d\tau \\
&\quad + \frac{p_2 S^*}{N^*}\int_0^A c_1 s_1 v_1(\tau)\bar{y}_1(\tau)d\tau + \frac{p_2 S^*}{N^*}\int_0^A c_2 s_2 v_2(\tau)\bar{y}_2(\tau)d\tau,
\end{aligned} \tag{3.30}$$

where

$$\bar{n} = \bar{x} + \int_0^A \bar{y}_1(\tau)d\tau + \int_0^A \bar{y}_2(\tau)d\tau.$$

Solving the differential equations (3.27) and (3.29), we obtain

$$\bar{y}_1(\tau) = \bar{y}_1(0)e^{-\lambda\tau}e^{-\int_0^\tau m(v_1(s))ds} \quad \textit{and} \quad \bar{y}_2(\tau) = \bar{y}_2(0)e^{-\lambda\tau}e^{-\int_0^\tau m(v_2(s))ds}.$$

From equations (3.26), (3.28) and (3.30), we have

$$\bar{y}_j(0) = -p_j(\lambda + m_0)\bar{x}, \quad j = 1, 2. \tag{3.31}$$

Using the definitions of $\bar{n}$, $\bar{y}_1(\tau)$, $\bar{y}_2(\tau)$, $\bar{y}_j(0)$, and setting $\alpha_j = \int_0^A c_j s_j v_j(\tau) i_j^*(\tau)d\tau$, equation (3.26) becomes

$$\begin{aligned}
&(\lambda + m_0)\bar{x} \\
&= -\frac{\bar{x}\alpha_1}{N^*} + \frac{S^*}{N^*}\frac{\alpha_1}{N^*}\Big(\bar{x} + \sum_{j=1}^2 \bar{y}_j(0)\int_0^A e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau\Big) \\
&\quad - \frac{\bar{x}\alpha_2}{N^*} + \frac{S^*}{N^*}\frac{\alpha_2}{N^*}\Big(\bar{x} + \sum_{j=1}^2 \bar{y}_j(0)\int_0^A e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau\Big) \\
&\quad - \frac{S^*}{N^*}\sum_{j=1}^2 \bar{y}_j(0)\int_0^A c_j s_j v_j(\tau)e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau \\
&= \frac{(\alpha_1+\alpha_2)\bar{x}}{N^*}\big(\frac{S^*}{N^*} - 1\big) + (\lambda + m_0)\bar{x}\frac{S^*}{N^*}\sum_{j=1}^2 p_j\int_0^A c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau \\
&\quad - \frac{(\alpha_1+\alpha_2)}{N^*}\frac{S^*}{N^*}(\lambda + m_0)\bar{x}\sum_{j=1}^2 p_j\int_0^A e^{-\lambda\tau}e^{-\int_0^\tau m(v_j(s))ds}d\tau,
\end{aligned} \tag{3.32}$$

because $\bar{y}_j(0)$ in defined in equation (3.31). Dividing both sides of equation (3.32) by $(\lambda + m_0)\bar{x}$, and substituting $\frac{S^*}{N^*} = \frac{1}{\mathcal{R}_0}$, we obtain the characteristic equation

$$1 = \frac{\alpha_1 + \alpha_2}{N^*\mathcal{R}_0}\Big(\frac{1-\mathcal{R}_0}{\lambda + m_0} - \sum_{j=1}^{2} p_j\Gamma_j(\lambda)\Big) + \frac{1}{\mathcal{R}_0}\sum_{j=1}^{2} p_j \int_0^A c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau, \quad (3.33)$$

where

$$\Gamma_j(\lambda) = \int_0^A e^{-\lambda\tau}\pi_j(\tau)d\tau \quad \text{and} \quad \pi_j(\tau) = e^{-\int_0^\tau m(v_j(s))ds}.$$

Now, using the mortality function, $m(v_j(\tau)) = m_0 + \mu_j v_j(\tau)$, and integration by parts, the term

$$\begin{aligned}
&\sum_{j=1}^{2} p_j \int_0^A c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau \\
&= \sum_{j=1}^{2} \frac{p_j c_j s_j}{\mu_j} \int_0^A \mu_j v_j(\tau)e^{-(\lambda+m_0)\tau}e^{-\int_0^\tau \mu_j v_j(s)ds}d\tau \qquad (3.34)\\
&= \sum_{j=1}^{2} \frac{p_j c_j s_j}{\mu_j}\big(1 - e^{-\lambda A}\pi_j(A) - (\lambda + m_0)\Gamma_j(\lambda)\big).
\end{aligned}$$

Thus, if $\lambda = 0$ in equation (3.34) and $\mathcal{R}_0 > 1$, then

$$1 < \mathcal{R}_0 = \sum_{j=1}^{2} \frac{p_j c_j s_j}{\mu_j}(1 - \pi_j(A)) - m_0 \sum_{j=1}^{2} \frac{p_j c_j s_j}{\mu_j}\Gamma_j(0).$$

Whence,

$$1 < \sum_{j=1}^{2} \frac{p_j c_j s_j}{\mu_j} \le \max\Big\{\frac{c_1 s_1}{\mu_1}, \frac{c_2 s_2}{\mu_2}\Big\}$$

due to the convex combination of $\frac{c_1 s_1}{\mu_1}$ and $\frac{c_2 s_2}{\mu_2}$. Now, using equation (3.34), equation (3.33) becomes

$$\begin{aligned}
&1 + \frac{\alpha_1 + \alpha_2}{N^*(\lambda + m_0)} \\
&= \frac{1}{\mathcal{R}_0}\frac{\alpha_1 + \alpha_2}{N^*(\lambda + m_0)} + \frac{1}{\mathcal{R}_0}\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau \\
&\quad - \frac{\alpha_1 + \alpha_2}{N^*\mathcal{R}_0}\frac{1}{\lambda + m_0}\frac{\mu_1}{c_1 s_1}\sum_{j=1}^{2}\frac{p_j c_j s_j}{\mu_j}(1 - e^{-\lambda A}\pi_j(A)) \\
&\quad + \frac{\alpha_1 + \alpha_2}{N^*\mathcal{R}_0}\frac{p_2 c_2 s_2}{\mu_2}\frac{\mu_1}{c_1 s_1}\Gamma_2(\lambda) - \frac{\alpha_1 + \alpha_2}{N^*\mathcal{R}_0}p_2\Gamma_2(\lambda) \\
&\quad + \frac{\alpha_1 + \alpha_2}{N^*\mathcal{R}_0}\frac{\mu_1}{c_1 s_1}\frac{1}{\lambda + m_0}\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau \\
&= \frac{1}{\mathcal{R}_0}\Big(1 + \frac{\alpha_1 + \alpha_2}{N^*(\lambda + m_0)}\frac{\mu_1}{c_1 s_1}\Big)\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau
\end{aligned}$$

$$+\frac{1}{\mathcal{R}_0}\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\Big(1-\frac{\mu_1}{c_1s_1}\sum_{j=1}^{2}\frac{p_jc_js_j}{\mu_j}(1-e^{-\lambda A}\pi_j(A))\Big)$$
$$-\frac{\alpha_1+\alpha_2}{N^*\mathcal{R}_0}\Big(1-\frac{c_2s_2}{\mu_2}\frac{\mu_1}{c_1s_1}\Big)p_2\Gamma_2(\lambda). \tag{3.35}$$

This gives

$$\begin{aligned}
&\frac{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}}{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\frac{\mu_1}{c_1s_1}}\\
&=\frac{1}{\mathcal{R}_0}\sum_{j=1}^{2}\int_0^A p_jc_js_jv_j(\tau)e^{-\lambda\tau}\pi_j(\tau)d\tau\\
&\quad+\frac{\frac{1}{\mathcal{R}_0}\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}}{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\frac{\mu_1}{c_1s_1}}\Big(1-\frac{\mu_1}{c_1s_1}\sum_{j=1}^{2}\frac{p_jc_js_j}{\mu_j}+\frac{\mu_1}{c_1s_1}\sum_{j=1}^{2}\frac{p_jc_js_j}{\mu_j}e^{-\lambda A}\pi_j(A)\Big)\\
&\quad-\frac{\frac{\alpha_1+\alpha_2}{N^*\mathcal{R}_0}}{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\frac{\mu_1}{c_1s_1}}\Big(1-\frac{c_2s_2}{\mu_2}\frac{\mu_1}{c_1s_1}\Big)p_2\Gamma_2(\lambda)=:\mathcal{L}(\lambda).
\end{aligned} \tag{3.36}$$

Now, if $\frac{c_1s_1}{\mu_1}=\frac{c_2s_2}{\mu_2}$, we obtain $1-\frac{c_2s_2}{\mu_2}\frac{\mu_1}{c_1s_1}=0$ and $1-\frac{\mu_1}{c_1s_1}\sum_{j=1}^{2}\frac{p_jc_js_j}{\mu_j}=0$. Thus, if $\Re(\lambda)>0$, then the left-hand side of equation (3.36) gives

$$\Big|\frac{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}}{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\frac{\mu_1}{c_1s_1}}\Big|>1 \tag{3.37}$$

and the corresponding right-hand side gives

$$\begin{aligned}
|\mathcal{L}(\lambda)|&\le\frac{1}{\mathcal{R}_0}\sum_{j=1}^{2}\int_0^A p_jc_js_jv_j(\tau)e^{-\Re(\lambda)\tau}\pi_j(\tau)d\tau\\
&\quad+\frac{1}{\mathcal{R}_0}\Big|\frac{\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}}{1+\frac{\alpha_1+\alpha_2}{N^*(\lambda+m_0)}\frac{\mu_1}{c_1s_1}}\Big|e^{-(\Re(\lambda)+m_0)A}.
\end{aligned}$$

Thus, $|\mathcal{L}(\lambda)|<1$ if A is sufficiently large. Thus, the case $\Re(\lambda)>0$ gives a contradiction. Next, if $\Re(\lambda)=0$ ($a=0$), we multiply both sides of the characteristic equation (3.35) by m_0+ib. This gives

$$\begin{aligned}
&\frac{\alpha_1+\alpha_2}{N^*}+m_0+ib\\
&=\frac{1}{\mathcal{R}_0}\Big(\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1s_1}+m_0+ib\Big)\sum_{j=1}^{2}\int_0^A p_jc_js_jv_j(\tau)e^{-ib\tau}\pi_j(\tau)d\tau\\
&\quad+\frac{1}{\mathcal{R}_0}\frac{\alpha_1+\alpha_2}{N^*}\Big(1-\frac{\mu_1}{c_1s_1}\sum_{j=1}^{2}\frac{p_jc_js_j}{\mu_j}(1-e^{-ibA}\pi_j(A))\Big)\\
&\quad-\frac{(m_0+ib)(\alpha_1+\alpha_2)}{N^*\mathcal{R}_0}\Big(1-\frac{c_2s_2}{\mu_2}\frac{\mu_1}{c_1s_1}\Big)p_2\Gamma_2(\lambda).
\end{aligned} \tag{3.38}$$

Equating the imaginary parts of equation (3.38), we obtain

$$\begin{aligned}
&b\Big(\mathcal{R}_0 - \sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)\cos(b\tau)\pi_j(\tau)d\tau\Big)\\
&= -\Big(\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1} + m_0\Big)\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)\sin(b\tau)\pi_j(\tau)d\tau\\
&\quad - \frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1}\sin(bA)\sum_{j=1}^{2}\frac{p_j c_j s_j}{\mu_j}\pi_j(A)\\
&\quad - \frac{b(\alpha_1+\alpha_2)}{N^*}\Big(1 - \frac{c_2 s_2}{\mu_2}\frac{\mu_1}{c_1 s_1}\Big)p_2\Gamma_2(\lambda)
\end{aligned} \tag{3.39}$$

Now, using the expression for the basic reproduction number (3.15), we have

$$\begin{aligned}
&\mathcal{R}_0 - \sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)\cos(b\tau)\pi_j(\tau)d\tau\\
&= \sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)(1-\cos(b\tau))\pi_j(\tau)d\tau\\
&= 2\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)\sin^2\big(\frac{b\tau}{2}\big)\pi_j(\tau)d\tau\\
&> 2\sum_{j=1}^{2} p_j c_j s_j \varepsilon_j'\pi_j(\alpha_2)\int_{\alpha_1}^{\alpha_2}\sin^2\big(\frac{b\tau}{2}\big)d\tau\\
&= \tilde{K}_2\pi(\alpha_2) > 0,
\end{aligned}$$

where ε_j' is the lower bound on $v_j(\tau)$ for $\tau \in [0, A]$ and $(\alpha_1, \alpha_2) \subset [0, A]$. Now, choose B^* such that

$$\begin{aligned}
&B^*\tilde{K}_2\pi(\alpha_2)\\
&> \Big(\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1} + m_0\Big)\sum_{j=1}^{2}\int_0^A p_j c_j s_j v_j(\tau)\pi_j(\tau)d\tau\\
&\quad + \frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1}\sum_{j=1}^{2}\frac{p_j c_j s_j}{\mu_j}\pi_j(A) + \frac{b(\alpha_1+\alpha_2)}{N^*}\Big(1 - \frac{c_2 s_2}{\mu_2}\frac{\mu_1}{c_1 s_1}\Big)p_2\Gamma_2(\lambda).
\end{aligned}$$

Then, for $b > B^*$, equation (3.39) is untenable. For $b < B^*$, the left-hand side of equation (3.36) gives

$$\Big|\frac{\frac{\alpha_1+\alpha_2}{N^*} + m_0 + ib}{\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1} + m_0 + ib}\Big| > \frac{\sqrt{(\frac{\alpha_1+\alpha_2}{N^*} + m_0)^2 + B^{*2}}}{\sqrt{(\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1} + m_0)^2 + B^{*2}}} > 1,$$

and the right-hand side of equation (3.36), with $\frac{c_1 s_1}{\mu_1} = \frac{c_2 s_2}{\mu_2}$ and $\Re(\lambda) = 0$ gives

$$|\mathcal{L}(\lambda)| \le 1 + \frac{\alpha_1+\alpha_2}{N^*\mathcal{R}_0}\frac{\sum_{j=1}^{2} p_j\pi_j(A)}{|\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1} + m_0 + ib|}$$

$$\leq 1 + \frac{\alpha_1+\alpha_2}{N^*\mathcal{R}_0}\frac{e^{-m_0 A}}{\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1}+m_0}$$
$$< \frac{\sqrt{(\frac{\alpha_1+\alpha_2}{N^*}+m_0)^2+B^{*2}}}{\sqrt{(\frac{\alpha_1+\alpha_2}{N^*}\frac{\mu_1}{c_1 s_1}+m_0)^2+B^{*2}}},$$

if A is sufficiently large. The case $\Re(\lambda) = 0$ is also a contradiction. Thus, the real parts of λ are non-positive, and hence, the endemic equilibrium is locally asymptotically stable if $\mathcal{R}_0 > 1$, A is sufficiently large and $\frac{c_1 s_1}{\mu_1} = \frac{c_2 s_2}{\mu_2}$. □

Remark 3.7. If $\frac{c_1 s_1}{\mu_1} \neq \frac{c_2 s_2}{\mu_2}$, one can use a numerical procedure to compute the basic reproduction number $\mathcal{R}_0$, using parameter values for HIV. If $\mathcal{R}_0 > 1$, then the characteristic equation (3.36) is solved numerically for λ (see Castillo-Chavez et al. [10]). Using different values of the c_j, s_j and μ_j the nature of roots of the equation (3.36) with largest real part may give an insight into the local stability of the endemic equilibrium for these parameter regimes.

4. Optimal control formulation and analysis

Optimal control of first-order PDEs coming from age-structured models requires more analysis for justification than optimal control of parabolic PDE or differential equations. There has been only a small amount of work on specific applications of optimal control to age-structured equations. Brokate [8] developed maximum principles for an optimal harvesting problem and a problem of optimal birth control. Barbu and Iannelli [6, 7] considered and optimal control problem for a Gurtin-MacCamy [22, 38] type system, describing the evolution of an age-structured population. Anita [3, 4] investigated an optimal control problem for a nonlinear age-dependent population dynamics. Murphy and Smith [33] studied the optimal harvesting of an age-structured population, where the McKendrick model of population dynamics was used. These authors considered age-structured population models for a single population. Fister and Lenhart [18], on the other hand, considered optimal harvesting control for a competitive age-structured model, comprising two first-order partial differential equations. Also, Fister and Lenhart [17] investigated an optimal harvesting control in a predator-prey model in which the prey population is represented by a first-order partial differential equation with age-structure and the predator is represented by an ordinary differential equation in time. Numfor et al [34] considered optimal control in coupled within-host and between-host models. The within-host model is a system of ODEs and the between-host model is a coupled system of ordinary and first-order PDEs. A key tool for the existence and uniqueness of optimal solution is Ekeland's variational principle [13].

In our multi-group coupled within-host and between-host model and in order to curtail the proliferation of free virus at the within-host level, we introduce two control functions u_1 and u_2, which delineate transmission and virion production suppressing drugs, respectively. This leads to the following multi-group within-host model

$$\frac{dx_j}{d\tau} = r - \beta_j(1-u_1(\tau))v_j(\tau)x_j(\tau) - \mu x_j(\tau) \tag{4.1}$$
$$\frac{dy_j}{d\tau} = \beta_j(1-u_1(\tau))v_j(\tau)x_j(\tau) - d_j y_j(\tau), \quad j = 1,2 \tag{4.2}$$

$$\frac{dv_j}{d\tau} = \gamma_j(1-u_2(\tau))d_j y_j(\tau) - (\delta_j + s_j)v_j(\tau) - \hat{\beta}_j(1-u_1(\tau))v_j(\tau)x_j(\tau), \quad (4.3)$$

We develop Lipschitz properties for the solutions to the state system in terms of controls. These properties will be used in proving the existence of sensitivities, and the existence and uniqueness of optimal control pair.

Theorem 4.1 (Lipschitz Property)**.** *The mapping*

$$(u_1,u_2) \to (x_1,x_2,y_1,y_2,v_1,v_2,S,i_1,i_2) = (x_1,x_2,y_1,y_2,v_1,v_2,S,i_1,i_2)(u_1,u_2)$$

is Lipschitz in the following ways: (i)

$$\sum_{j=1}^{2}\int_\Omega (|x_j-\bar{x}_j| + |y_j-\bar{y}_j| + |v_j-\bar{v}_j|)d\tau + \int_0^T |S-\bar{S}|dt + \sum_{j=1}^{2}\int_Q |i_j-\bar{i}_j|d\tau dt$$
$$\le C_{A,T}\int_\Omega (|u_1-\bar{u}_1| + |u_2-\bar{u}_2|)d\tau$$

(ii)

$$\begin{aligned}&\|S-\bar{S}\|_{L^\infty(0,T)} + \sum_{j=1}^{2}\Big(\|x_j-\bar{x}_j\|_{L^\infty(\Omega)} + \|y_j-\bar{y}_j\|_{L^\infty(\Omega)}\\ &+ \|v_j-\bar{v}_j\|_{L^\infty(\Omega)} + \|i_j-\bar{i}_j\|_{L^\infty(Q)}\Big)\\ &\le \hat{C}_{A,T}(\|u_1-\bar{u}_1\|_{L^\infty(\Omega)} + \|u_2-\bar{u}_2\|_{L^\infty(\Omega)}),\end{aligned}$$

where $\Omega = (0,A)$ *and* $Q = \Omega\times(0,T)$.

The proof of the above theorem is the proof in Numfor et al. [34, Theorem 3.2].

4.1. **The optimality system.** In this subsection, we derive a sensitivity system, an adjoint system and a control characterization. To derive a characterization of an optimal control, we define an objective functional, J, for our problem, where our objective is to minimize free virus, population of infectious individuals and the cost of implementing the control. Thus, we use the objective functional

$$\begin{aligned}J(u_1,u_2) = &\int_0^T\int_0^A (A_1 i_1(\tau,t)v_1(\tau) + i_1(\tau,t)(A_2u_1(\tau)+A_3u_2(\tau)))d\tau dt\\ &+ \int_0^T\int_0^A (A_4 i_2(\tau,t)v_2(\tau) + i_2(\tau,t)(A_2u_1(\tau)+A_3u_2(\tau)))d\tau dt\\ &+ \int_0^A (B_1u_1(\tau)^2 + B_2u_2(\tau)^2)d\tau,\end{aligned} \quad (4.4)$$

where A_1, A_2, A_3, A_4, B_1 and B_2 are positive constants that balance the relative importance for the terms in J. The term $\int_0^T\int_0^A (A_1i_1(\tau,t)v_1(\tau) + A_4i_2(\tau,t)v_2(\tau))d\tau dt$ in the objective functional gives the total of infected individuals in the population over the time period T and age-since-infection A to be minimized. The terms $i_1(\tau,t)u_1(\tau)$ and $i_2(\tau,t)u_1(\tau)$ represent the number of infected individuals treated with the transmission suppressing drug respectively, and A_2 is the cost per individual treated with this drug. Thus,

$$\int_0^T\int_0^A (A_2i_1(\tau,t)u_1(\tau) + A_2i_2(\tau,t)u_1(\tau))d\tau dt + \int_0^A B_1u_1^2(\tau)d\tau$$

gives the cost of implementing the control with the transmission suppressing drug for all infected individuals of age-since-infection, A. Here, we assume a nonlinear cost for treatment and chose the quadratic cost for illustration. By analogy, we define other terms in the objective functional.

The optimal control formulation for our problem is: Find $(u_1^*, u_2^*) \in \mathcal{U}$ such that

$$J(u_1^*, u_2^*) = \min_{(u_1,u_2)\in\mathcal{U}} J(u_1, u_2),$$

where the set of all admissible controls is

$$\mathcal{U} = \{(u_1, u_2) \in L^\infty(0, A) \times L^\infty(0, A) | u_1 : (0, A) \to [0, \tilde{u}_1],\ u_2 : (0, A) \to [0, \tilde{u}_2]\}.$$

The upper bounds on the controls give the efficacy of the transmission and virion production suppressing drugs while the lower bounds, $u_1 = 0$ and $u_2 = 0$, represent the case where there is no inhibition of transmission and virion production.

We take the Gâteaux derivatives of J with respect to controls $(u_1, u_2) \in \mathcal{U}$. Since the objective functional is defined in term of the states, we start by finding the derivatives of the control-to-state map. These derivatives are called sensitivities.

Theorem 4.2 (Sensitivities)**.** *The map*

$$(u_1, u_2) \to (x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2) = (x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2)(u_1, u_2)$$

is differentiable in the following sense:

$$\frac{\Phi(u_1 + \varepsilon l_1, u_2 + \varepsilon l_2) - \Phi(u_1, u_2)}{\varepsilon} \to (\psi_1, \psi_2, \varphi_1, \varphi_2, \phi_1, \phi_2, \theta, \omega_1, \omega_2)$$

in $(L^\infty(\Omega))^6 \times L^\infty(0, T) \times (L^\infty(0, T; L^1(\Omega)))^2$, *as* $\varepsilon \to 0$ *with* $(u_1 + \varepsilon l_1, u_2 + \varepsilon l_2)$, $(u_1, u_2) \in \mathcal{U}$ *and* $l_1, l_2 \in L^\infty(\Omega)$, *where* $\Phi = (x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2)$. *Furthermore, for* $j = 1, 2$, *the sensitivity functions satisfy*

$$\frac{d\psi_j}{d\tau} = -(\beta_j(1-u_1)v_j + \mu)\psi_j - \beta_j(1-u_1)x_j\phi_j + \beta_j l_1 v_j x_j \tag{4.5}$$

$$\frac{d\varphi_j}{d\tau} = \beta_j(1-u_1)v_j\psi_j - d_j\varphi_j + \beta_j(1-u_1)x_j\phi_j - \beta_j l_1 v_j x_j \tag{4.6}$$

$$\begin{aligned}\frac{d\phi_j}{d\tau} &= -\hat{\beta}_j(1-u_1)v_j\psi_j + \gamma_j(1-u_2)d_j\varphi_j - (\delta_j + s_j + \hat{\beta}_j(1-u_1)x_j)\phi_j \\ &\quad + \hat{\beta}_j l_1 v_j x_j - \gamma_j d_j l_2 y_j\end{aligned} \tag{4.7}$$

$$\begin{aligned}\frac{d\theta}{dt} &= -m_0\theta(t) - \frac{1}{N(t)}\left(1 - \frac{S(t)}{N(t)}\right)\theta(t)\sum_{k=1}^{2} c_k s_k \int_\Omega i_k(\tau, t)v_k(\tau)d\tau \\ &\quad - \frac{S(t)}{N(t)}\sum_{k=1}^{2} c_k s_k \int_\Omega [v_k(\tau)\omega_k(\tau, t) + i_k(\tau, t)\phi_k(\tau)]d\tau \\ &\quad + \frac{S(t)}{N(t)^2}\int_\Omega (\omega_1(h, t) + \omega_2(h, t))dh \sum_{k=1}^{2} c_k s_k \int_\Omega i_k(\tau, t)v_k(\tau)d\tau \quad \text{in } (0, T)\end{aligned} \tag{4.8}$$

$$\frac{\partial\omega_j}{\partial t} + \frac{\partial\omega_j}{\partial\tau} = -m(v_j(\tau))\omega_j(\tau, t) - \mu_j\phi_j(\tau)i_j(\tau, t) \quad \text{in } \Omega \times (0, T) \tag{4.9}$$

with initial and boundary conditions

$$\psi_j(0) = 0, \quad \varphi_j(0) = 0, \quad \phi_j(0) = 0, \quad \theta(0) = 0, \quad \omega_j(\tau, 0) = 0, \tag{4.10}$$

for $\tau \in \Omega = (0, A)$, *and*

$$\begin{aligned}\omega_j(0,t) &= \frac{p_j}{N(t)}\Big(1-\frac{S(t)}{N(t)}\Big)\theta(t)\sum_{k=1}^{2} c_k s_k \int_\Omega i_k(\tau,t)v_k(\tau)d\tau \\ &\quad + p_j\frac{S(t)}{N(t)}\sum_{k=1}^{2} c_k s_k \int_\Omega [v_k(\tau)\omega_k(\tau,t) + i_k(\tau,t)\phi_k(\tau)]d\tau \\ &\quad - p_j\frac{S(t)}{N(t)^2}\int_\Omega (\omega_1(h,t)+\omega_2(h,t))dh\sum_{k=1}^{2} c_k s_k \int_\Omega i_k(\tau,t)v_k(\tau)d\tau.\end{aligned} \tag{4.11}$$

Proof. Since the map $(u_1, u_2) \to (x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2)$ is Lipschitz in L^∞, we have the existence of Gâteaux derivatives $\psi_1, \psi_2, \varphi_1, \varphi_2, \phi_1, \phi_2, \theta, \omega_1, \omega_2$ by Barbu [7], Fister et al [17, 18] and Numfor et al [34]. Passing to the limit in the representation of the difference quotients in state functions, the sensitivity functions $\psi_1, \psi_2, \varphi_1, \varphi_2, \phi_1, \phi_2, \theta, \omega_1, \omega_2$ satisfy system (4.5)–(4.11). □

From the sensitivity equations in Theorem 4.2, we introduce three sensitivity operators, $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$, satisfying the the sensitivity equations:

$$\mathcal{L}_1\begin{bmatrix}\psi_1\\ \psi_2\\ \varphi_1\\ \varphi_2\\ \phi_1\\ \phi_2\end{bmatrix} = \begin{bmatrix}\beta_1 l_1 v_1 x_1\\ \beta_2 l_1 v_2 x_2\\ -\beta_1 l_1 v_1 x_1\\ -\beta_2 l_1 v_2 x_2\\ \hat{\beta}_1 l_1 v_1 x_1 - \gamma_1 d_1 l_2 y_1\\ \hat{\beta}_2 l_1 v_2 x_2 - \gamma_2 d_2 l_2 y_2\end{bmatrix}, \quad \mathcal{L}_2\theta = 0, \quad \mathcal{L}_3\begin{bmatrix}\omega_1\\ \omega_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}. \tag{4.12}$$

Using the sensitivity system, we derive the adjoint system. Thus, if λ_1,λ_2, ξ_1, ξ_2, η_1, η_2, p, q_1 and q_2 are adjoint functions, then we find adjoint operators $\mathcal{L}_j^*$, for $j = 1, 2, 3$ satisfying

$$\begin{gathered}\mathcal{L}_1^*\begin{bmatrix}\lambda_1\\ \lambda_2\\ \xi_1\\ \xi_2\\ \eta_1\\ \eta_2\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\\ 0\\ A_1\int_0^T i_1(\tau,t)dt\\ A_4\int_0^T i_2(\tau,t)dt\end{bmatrix}, \\ \mathcal{L}_2^* p = 0, \quad \mathcal{L}_3^*\begin{bmatrix}q_1\\ q_2\end{bmatrix} = \begin{bmatrix}A_1v_1 + A_2u_1 + A_3u_2\\ A_4v_2 + A_2u_1 + A_3u_2\end{bmatrix}.\end{gathered} \tag{4.13}$$

The right-hand side of the adjoint operators (4.13) are obtained by differentiating the integrand of the objective functional (4.4) with respect to each state variable. The transversality conditions associated with the adjoint variables are:

$$\lambda_j(A) = 0, \quad \xi_j(A) = 0, \quad \eta_j(A) = 0, \quad p(T) = 0 \tag{4.14}$$

$$q_j(\tau, T) = 0, \quad \text{for } \tau \in (0, A) \tag{4.15}$$

$$q_j(A, t) = 0, \quad \text{for } t \in (0, T) \text{ and } j = 1, 2. \tag{4.16}$$

From the sensitivity system in Theorem 4.2 and using a relationship between the sensitivity and adjoint operators in terms of their inner product in L^2, we use integration by parts to throw the derivatives in the differential operators in the

sensitivity functions ψ_j, φ_j, ϕ_j, θ, and ω_j onto the adjoint functions λ_j, ξ_j, η_j, p and q_j to form the adjoint operators.

Using the relationship between the sensitivity and adjoint operators, we have the following system of adjoint equations corresponding to controls (u_1, u_2), and states $(x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2) = (x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2)(u_1, u_2)$:

$$-\frac{d\lambda_1}{d\tau} = -(\beta_1(1-u_1)v_1 + \mu)\lambda_1 + \beta_1(1-u_1)v_1\xi_1 - \hat{\beta}_1(1-u_1)v_1\eta_1 \tag{4.17}$$

$$-\frac{d\lambda_2}{d\tau} = -(\beta_2(1-u_1)v_2 + \mu)\lambda_2 + \beta_2(1-u_1)v_2\xi_2 - \hat{\beta}_2(1-u_1)v_2\eta_2 \tag{4.18}$$

$$-\frac{d\xi_1}{d\tau} = -d_1\xi_1 + d_1\gamma_1(1-u_2)\eta_1 \tag{4.19}$$

$$-\frac{d\xi_2}{d\tau} = -d_2\xi_2 + d_2\gamma_2(1-u_2)\eta_2 \tag{4.20}$$

$$\begin{aligned}
-\frac{d\eta_1}{d\tau} &= -\beta_1(1-u_1)x_1\lambda_1 + \beta_1(1-u_1)x_1\xi_1 - (\delta_1 + s_1 + \hat{\beta}_1(1-u_1)x_1)\eta_1 \\
&\quad - c_1s_1\int_0^T \frac{S(t)i_1(\tau,t)}{N(t)}p(t)dt - m'(v_1)\int_0^T i_1(\tau,t)q_1(\tau,t)dt \\
&\quad + c_1s_1\int_0^T (p_1q_1(0,t) + p_2q_2(0,t))\frac{S(t)i_1(\tau,t)}{N(t)}dt + A_1\int_0^T i_1(\tau,t)dt \\
-\frac{d\eta_2}{d\tau} &= -\beta_2(1-u_1)x_2\lambda_2 + \beta_2(1-u_1)x_2\xi_2 - (\delta_2 + s_2 + \hat{\beta}_2(1-u_1)x_2)\eta_2 \\
&\quad - c_2s_2\int_0^T \frac{S(t)i_2(\tau,t)}{N(t)}p(t)dt - m'(v_2)\int_0^T i_2(\tau,t)q_2(\tau,t)dt \\
&\quad + c_2s_2\int_0^T (p_1q_1(0,t) + p_2q_2(0,t))\frac{S(t)i_2(\tau,t)}{N(t)}dt + A_4\int_0^T i_2(\tau,t)dt
\end{aligned} \tag{4.21}$$

$$\begin{aligned}
-\frac{dp}{dt} &= -m_0p - \frac{p}{N}\big(1-\frac{S}{N}\big)\sum_{j=1}^{2} c_js_j\int_0^A v_j(\tau)i_j(\tau,t)d\tau \\
&\quad + \frac{p_1q_1(0,t) + p_2q_2(0,t)}{N}\big(1-\frac{S}{N}\big)\sum_{j=1}^{2} c_js_j\int_0^A i_j(\tau,t)v_j(\tau)d\tau
\end{aligned} \tag{4.22}$$

$$\begin{aligned}
&-\frac{\partial q_1}{\partial t} - \frac{\partial q_1}{\partial \tau} \\
&= -m(v_1)q_1 - c_1s_1(p(t) - p_1q_1(0,t) - p_2q_2(0,t))\frac{Sv_1}{N} \\
&\quad + (p(t) - p_1q_1(0,t) - p_2q_2(0,t))\frac{S}{N^2}\sum_{j=1}^{2} c_js_j\int_0^A i_j(\tau,t)v_j(\tau)d\tau \\
&\quad + A_1v_1 + A_2u_1 + A_3u_2
\end{aligned} \tag{4.23}$$

$$\begin{aligned}
&-\frac{\partial q_2}{\partial t} - \frac{\partial q_2}{\partial \tau} \\
&= -m(v_2)q_2 - c_2s_2(p(t) - p_1q_1(0,t) - p_2q_2(0,t))\frac{Sv_2}{N}
\end{aligned}$$

$$+ (p(t) - p_1q_1(0,t) - p_2q_2(0,t))\frac{S}{N^2}\sum_{j=1}^{2} c_j s_j \int_0^A i_j(\tau,t)v_j(\tau)d\tau$$
$$+ A_4v_2 + A_2u_1 + A_3u_2, \tag{4.24}$$

with final time conditions given in equations (4.14)–(4.16).

The weak solution to our problem is characterized in Theorem 4.3. This solution is used in characterizing the solution to the adjoint system which satisfies a Lipschitz property analogous to Theorem 4.1. This property will be used in proving existence and uniqueness of an optimal control pair.

Theorem 4.3. *The weak solution of the adjoint system satisfies*

$$\begin{aligned}
0 &= \sum_{j=1}^{2}\int_0^A (\lambda_j\alpha_j + \xi_j\tilde{\alpha}_j + \eta_j\hat{\alpha}_j)d\tau \\
&\quad - \int_0^T\int_0^A (A_1g_1(\tau)i_1(\tau,t) + A_4g_2(\tau)i_2(\tau,t))d\tau dt \\
&\quad - \int_0^T\int_0^A (A_1v_1(\tau) + A_2u_1(\tau) + A_3u_2(\tau))n_1(\tau,t)d\tau dt \\
&\quad - \int_0^T\int_0^A (A_4v_2(\tau) + A_2u_1(\tau) + A_3u_2(\tau))n_2(\tau,t)d\tau dt,
\end{aligned}$$

where for $j = 1, 2$, the functions α_j, $\tilde{\alpha}_j$, $\hat{\alpha}_j$ in $L^\infty(0,A)$ are obtained from test functions z_j, f_j and g_j, and r and n_j satisfy equations (4.22)–(4.24) such that

$$\frac{dz_j}{d\tau} + (\beta_j(1-u_1)v_j + \mu)z_j + \beta_j(1-u_1)x_jg_j = \alpha_j$$
$$\frac{df_j}{d\tau} - \beta_j(1-u_1)v_jz_j + d_jf_j - \beta_j(1-u_1)x_jg_j = -\tilde{\alpha}_j,$$
$$\frac{dg_j}{d\tau} + \hat{\beta}_j(1-u_1)v_jz_j - \gamma_j(1-u_2)d_jf_j + (\delta_j + s_j + \hat{\beta}_j(1-u_1)x_j)z_j = \hat{\alpha}_j,$$
$$\frac{dr}{dt} + m_0r(t) + \frac{1}{N(t)}\big(1 - \frac{S(t)}{N(t)}\big)r(t)\sum_{k=1}^{2} c_ks_k\int_\Omega i_k(\tau,t)v_k(\tau)d\tau$$
$$+ \frac{S(t)}{N(t)}\sum_{k=1}^{2} c_ks_k\int_\Omega [v_k(\tau)n_k(\tau,t) + i_k(\tau,t)z_k(\tau)]d\tau$$
$$- \frac{S(t)}{N(t)^2}\sum_{k=1}^{2} c_ks_k\int_\Omega i_k(\tau,t)v_k(\tau)\int_\Omega (n_1(h,t) + n_2(h,t))dhd\tau = 0 \quad \text{in } (0,T),$$
$$\frac{\partial n_j}{\partial t} + \frac{\partial n_j}{\partial \tau} + m(v_j(\tau))n_j(\tau,t) + m'(v_j(\tau))z_j(\tau)i_j(\tau,t) = 0 \quad \text{in } \Omega\times(0,T),$$

with boundary and initial conditions

$$\begin{aligned}
n_j(0,t) &= \frac{p_j}{N(t)}\left(1 - \frac{S(t)}{N(t)}\right)r(t)\sum_{k=1}^{2} c_ks_k\int_\Omega i_k(\tau,t)v_k(\tau)d\tau \\
&\quad + p_j\frac{S(t)}{N(t)}\sum_{k=1}^{2} c_ks_k\int_\Omega [v_k(\tau)n_k(\tau,t) + i_k(\tau,t)z_k(\tau)]d\tau
\end{aligned}$$

$$- p_j \frac{S(t)}{N(t)^2} \sum_{k=1}^{2} c_k s_k \int_\Omega i_k(\tau,t) v_k(\tau) \int_\Omega (n_1(h,t) + n_2(h,t)) dh d\tau,$$

and

$$z_j(0) = 0, \quad f_j(0) = 0, \quad g_j(0) = 0, \quad r(0) = 0, \quad n_j(\tau,0) = 0, \quad \text{for } \tau \in \Omega.$$

The proof of the above theorem follows from the sensitivity equations and adjoint system, with $\alpha_j = \beta_j l_1 v_j x_j$, $\tilde{\alpha}_j = \beta_j l_1 v_j x_j$ and $\hat{\alpha}_j = \hat{\beta}_j l_1 v_j x_j - \gamma_j d_j l_2 y_j$.

Theorem 4.4. *For $(u_1, u_2) \in \mathcal{U}$, the adjoint system (4.17)–(4.24) has a weak solution $(\lambda_1, \lambda_2, \xi_1, \xi_2, \eta_1, \eta_2, p, q_1, q_2)$ in $(L^\infty(0,A))^6 \times L^\infty(0,T) \times (L^\infty(0,T,L^1(0,A)))^2$ such that*

$$\begin{aligned}
&\sum_{j=1}^{2} \big(\|\lambda_j - \bar{\lambda}_j\|_{L^\infty(\Omega)} + \|\xi_j - \bar{\xi}_j\|_{L^\infty(\Omega)} + \|\eta_j - \bar{\eta}_j\|_{L^\infty(\Omega)} \big) + \|p - \bar{p}\|_{L^\infty(0,T)} \\
&+ \sum_{j=1}^{2} \|q_j - \bar{q}_j\|_{L^\infty(Q)} \\
&\le \tilde{C}_{A,T} (\|u_1 - \bar{u}_1\|_{L^\infty(\Omega)} + \|u_2 - \bar{u}_2\|_{L^\infty(\Omega)}).
\end{aligned}$$

The proof of the above theorem follows the steps of Theorem 4.1, part (ii).

We characterize the optimal control pair (u_1^*, u_2^*) by differentiating the control-to-objective functional map. Since the solutions of first-order partial differential equations are less regular than the solutions of parabolic PDEs, the method used in characterizing optimal control of first-order PDEs is different from that of parabolic PDEs. We use the Ekeland's principle [3, 13] to characterize optimal control of first-order PDEs. To do this, we embed the objective functional J in the space $L^1(\Omega) \times L^1(Q)$ by defining

$$\mathcal{J}(u_1, u_2) = \begin{cases} J(u_1, u_2) & \text{if } (u_1, u_2) \in \mathcal{U} \\ +\infty & \text{if } (u_1, u_2) \notin \mathcal{U}; \end{cases} \tag{4.25}$$

see [6, 17, 18]. To characterize the optimal control pair, we differentiate the objective functional, $\mathcal{J}$, with respect to the controls. However, since the objective functional is a function of the state functions, we must differentiate the state functions with respect to the controls.

Theorem 4.5 (Characterization)**.** *If $(u_1^*, u_2^*) \in \mathcal{U}$ is an optimal control pair minimizing (4.25), and $(x_1^*, x_2^*, y_1^*, y_2^*, v_1^*, v_2^*, S^*, i_1^*, i_2^*)$ and $(\lambda_1, \lambda_2, \xi_1, \xi_2, \eta_1, \eta_2, p, q_1, q_2)$ are the corresponding state and adjoint solutions, respectively, then*

$$u_1^*(\tau) = \mathcal{H}_1\Big(\frac{a_1^*(\tau) + a_2^*(\tau) - A_2 \int_0^T (i_1^*(\tau,t) + i_2^*(\tau,t)) dt}{2B_1} \Big), \tag{4.26}$$

$$u_2^*(\tau) = \mathcal{H}_2\Big(\frac{a_3^*(\tau) - A_3 \int_0^T (i_1^*(\tau,t) + i_2^*(\tau,t)) dt}{2B_2} \Big) \quad \text{a.e. in } L^1(\Omega), \tag{4.27}$$

where

$$\begin{aligned}
a_1^*(\tau) &= \beta_1 v_1^*(\tau) x_1^*(\tau)(\xi_1(\tau) - \lambda_1(\tau)) - \hat{\beta}_1 v_1^*(\tau) x_1^*(\tau) \eta_1(\tau) \\
a_2^*(\tau) &= \beta_2 v_2^*(\tau) x_2^*(\tau)(\xi_2(\tau) - \lambda_2(\tau)) - \hat{\beta}_2 v_2^*(\tau) x_2^*(\tau) \eta_2(\tau) \\
a_3^*(\tau) &= \gamma_1 d_1 \eta_1(\tau) y_1^*(\tau) + \gamma_2 d_2 \eta_2(\tau) y_2^*(\tau),
\end{aligned} \tag{4.28}$$

and for $j = 1, 2$,

$$\mathcal{H}_j(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le \tilde{u}_j, \\ \tilde{u}_j, & x > \tilde{u}_j. \end{cases}$$

Proof. Since (u_1^*, u_2^*) is an optimal control pair and we seek to minimize our functional, we have

$$\begin{aligned}
0 &\le \lim_{\varepsilon \to 0^+} \frac{\mathcal{J}(u_1^* + \varepsilon l_1, u_2^* + \varepsilon l_2) - \mathcal{J}(u_1^*, u_2^*)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0^+} \int_0^T \int_0^A \Big(A_1 v_1^\varepsilon \big(\frac{i_1^\varepsilon - i_1^*}{\varepsilon}\big) + A_1 i_1^* \big(\frac{v_1^\varepsilon - v_1^*}{\varepsilon}\big) + \frac{A_2(i_1^\varepsilon u_1^\varepsilon - i_1^* u_1^*)}{\varepsilon} \Big) d\tau dt \\
&\quad + \lim_{\varepsilon \to 0^+} \int_0^T \int_0^A \Big(A_4 v_2^\varepsilon \big(\frac{i_2^\varepsilon - i_2^*}{\varepsilon}\big) + A_4 i_2^* \big(\frac{v_2^\varepsilon - v_2^*}{\varepsilon}\big) + \frac{A_2(i_2^\varepsilon u_1^\varepsilon - i_2^* u_1^*)}{\varepsilon} \Big) d\tau dt \\
&\quad + \lim_{\varepsilon \to 0^+} \int_0^T \int_0^A \Big(\frac{A_3\big(i_1^\varepsilon u_2^\varepsilon - i_1^* u_2^*\big)}{\varepsilon} + \frac{A_3(i_2^\varepsilon u_2^\varepsilon - i_2^* u_2^*)}{\varepsilon} \Big) d\tau dt \\
&\quad + \lim_{\varepsilon \to 0^+} \int_0^A \Big(\frac{B_1((u_1^\varepsilon)^2 - (u_1^*)^2)}{\varepsilon} + \frac{B_2((u_2^\varepsilon)^2 - (u_2^*)^2)}{\varepsilon} \Big) d\tau \\
&= \int_0^T \int_0^A [(A_1 v_1^* \omega_1 + A_1 i_1^* \phi_1 + (A_2 u_1^* + A_3 u_2^*)\omega_1] d\tau dt \\
&\quad + \int_0^T \int_0^A [(A_4 v_2^* \omega_2 + A_4 i_2^* \phi_2 + (A_2 u_1^* + A_3 u_2^*)\omega_2] d\tau dt \\
&\quad + \int_0^T \int_0^A (A_2 l_1 (i_1^* + i_2^*) + A_3 l_2 (i_1^* + i_2^*)) d\tau dt + 2 \int_0^A (B_1 l_1 u_1^* + B_2 l_2 u_2^*) d\tau \\
&= \int_0^A l_1 \Big(\beta_1 v_1^* x_1^* (\lambda_1 - \xi_1) + \hat{\beta}_1 v_1^* x_1^* \eta_1 + \beta_2 v_2^* x_2^* (\lambda_2 - \xi_2) + \hat{\beta}_2 v_2^* x_2^* \eta_2 \\
&\quad + 2B_1 u_1^* + A_2 \int_0^T (i_1^* + i_2^*) dt \Big) d\tau \\
&\quad + \int_0^A l_2 \Big(2B_2 u_2^* - \gamma_1 d_1 y_1^* \eta_1 - \gamma_2 d_2 y_2^* \eta_2 + A_3 \int_0^T (i_1^*(\tau, t) + i_2^*(\tau, t)) dt \Big) d\tau.
\end{aligned}$$

Considering cases on the sets $\{\tau \in \Omega | u_j^*(\tau) = 0\}$, $\{\tau \in \Omega | u_j^*(\tau) = \tilde{u}_j\}$ and $\{\tau \in \Omega | 0 < u_j^*(\tau) < \tilde{u}_j\}$, for $j = 1, 2$, and using standard arguments, we obtain the desired characterization given in equations (4.26) and (4.27). □

4.2. **Existence of optimal control pair.** Existence results are obtained via Ekeland's principle. In order to use Ekeland's principle, we prove that our objective functional is lower semi-continuous with respect to L^1 convergence. On the other hand, uniqueness of optimal control pair is established by using the Lipschitz properties of the state and adjoint solutions given in Theorems 4.1 and 4.4, respectively, as well as the minimizing sequence obtained from the Ekeland's principle.

Theorem 4.6 (Lower semi-continuity)**.** *The functional* $\mathcal{J} : L^1(\Omega) \times L^1(\Omega) \to (-\infty, +\infty]$ *is lower semi-continuous.*

Given a lower semi-continuous functional, $\mathcal{J}$, we have the following Ekeland's principle which guarantees the existence of minimizers of an approximate functional,

$\mathcal{J}_\varepsilon$: For $\varepsilon > 0$, there exist $(u_1^\varepsilon, u_2^\varepsilon) \in L^1(0,A) \times L^1(0,A)$ such that

$$\mathcal{J}(u_1^\varepsilon, u_2^\varepsilon) \le \inf_{(u_1,u_2)\in\mathcal{U}} \mathcal{J}(u_1,u_2) + \varepsilon$$
$$\mathcal{J}(u_1^\varepsilon, u_2^\varepsilon) = \min_{(u_1,u_2)\in\mathcal{U}} \mathcal{J}_\varepsilon(u_1,u_2),$$

where

$$\mathcal{J}_\varepsilon(u_1,u_2) = \mathcal{J}(u_1,u_2) + \sqrt{\varepsilon}(\|u_1^\varepsilon - u_1\|_{L^1(0,A)} + \|u_2^\varepsilon - u_2\|_{L^1(0,A)}).$$

Theorem 4.7. *If $(u_1^\varepsilon, u_2^\varepsilon)$ is an optimal control pair minimizing the approximate functional, $\mathcal{J}_\varepsilon$, then*

$$(u_1^\varepsilon(\tau), u_2^\varepsilon(\tau)) = \mathcal{H}\Big(\frac{e_1^\varepsilon(\tau) + e_2^\varepsilon(\tau) - A_2K^\varepsilon(\tau) - \sqrt{\varepsilon}\kappa_1^\varepsilon(\tau)}{2B_1}, \frac{e_3^\varepsilon(\tau) - A_3K^\varepsilon(\tau) - \sqrt{\varepsilon}\kappa_2^\varepsilon(\tau)}{2B_2}\Big),$$

where

$$\begin{aligned} e_1^\varepsilon(\tau) &= \beta_1 v_1^\varepsilon(\tau)x_1^\varepsilon(\tau)(\xi_1^\varepsilon(\tau) - \lambda_1^\varepsilon(\tau)) - \hat{\beta}_1 v_1^\varepsilon(\tau)x_1^\varepsilon(\tau)\eta_1^\varepsilon(\tau) \\ e_2^\varepsilon(\tau) &= \beta_2 v_2^\varepsilon(\tau)x_2^\varepsilon(\tau)(\xi_2^\varepsilon(\tau) - \lambda_2^\varepsilon(\tau)) - \hat{\beta}_2 v_2^\varepsilon(\tau)x_2^\varepsilon(\tau)\eta_2^\varepsilon(\tau) \\ e_3^\varepsilon(\tau) &= \gamma_1 d_1\eta_1(\tau)y_1^\varepsilon(\tau) + \gamma_2 d_2 y_2^\varepsilon(\tau)\eta_2^\varepsilon(\tau) \\ K^\varepsilon(\tau) &= \int_0^T (i_1^\varepsilon(\tau,t) + i_2^\varepsilon(\tau,t))dt, \end{aligned} \tag{4.29}$$

and the functions $\kappa_1, \kappa_2 \in L^\infty(0,A)$, with $|\kappa_1(\tau)| = 1$ and $|\kappa_2(\tau)| = 1$, for all $\tau \in (0,A)$.

4.3. **Uniqueness of optimal control pair.** The uniqueness of an optimal control pair for the multi-group coupled within-host and between-host model is established using the Lipschitz properties for the state and adjoint functions in terms of the control functions in Theorems 4.1 and 4.4, and Ekelands variational principle.

Theorem 4.8. *If $\frac{\bar{C}_{A,T}}{2}(\frac{1}{B_1} + \frac{1}{B_2})$ is sufficiently small, then there exists a unique optimal control pair $(u_1^*, u_2^*) \in \mathcal{U}$ minimizing the objective functional $\mathcal{J}$.*

Proof. Let $\mathcal{H}(x,y) = (\mathcal{H}_1(x), \mathcal{H}_2(y))$ and define $L : \mathcal{U} \to \mathcal{U}$, such that

$$L(u_1,u_2) = \mathcal{H}\Big(\frac{a_1 + a_2 - A_2K(\tau)}{2B_1}, \frac{\gamma_1 d_1\eta_1 y_1 - A_3K(\tau)}{2B_2}\Big),$$

where a_j, $j = 1,2$ are defined in equation (4.28). Let $(x_1, x_2, y_1, y_2, v_1, v_2, S, i_1, i_2)$ and $(\lambda_1, \lambda_2, \xi_1, \xi_2, \eta_1, \eta_2, p, q_1, q_2)$ be state and adjoint solutions corresponding to the control pair (u_1, u_2). Then

$$\begin{aligned} &\|L(u_1,u_2) - L(\bar{u}_1,\bar{u}_2)\|_{L^\infty(0,A)\times L^\infty(0,A)} \\ &\equiv \|\mathcal{H}_1(u_1) - \mathcal{H}_1(\bar{u}_1)\|_{L^\infty(0,A)} + \|\mathcal{H}_2(u_2) - \mathcal{H}_2(\bar{u}_2)\|_{L^\infty(0,A)} \\ &\le \|\frac{e_1 + e_2 - A_2K(\tau)}{2B_1} - \frac{\bar{e}_1 + \bar{e}_2 - A_2\bar{K}(\tau)}{2B_1}\|_{L^\infty(0,A)} \\ &\quad + \|\frac{e_3 - A_3K(\tau)}{2B_2} - \frac{\bar{e}_3 - A_3\bar{K}(\tau)}{2B_2}\|_{L^\infty(0,A)} \\ &\le \frac{1}{2B_1}\|e_1 - \bar{e}_1\|_{L^\infty(0,A)} + \frac{1}{2B_1}\|e_2 - \bar{e}_2\|_{L^\infty(0,A)} + \frac{1}{2B_2}\|e_3 - \bar{e}_3\|_{L^\infty(0,A)} \\ &\quad + \frac{1}{2}\big(\frac{A_2}{B_1} + \frac{A_3}{B_2}\big)\|K - \bar{K}\|_{L^\infty(0,A)}, \end{aligned}$$

where for $j = 1, 2$,

$$\begin{aligned} e_j - \bar{e}_j &= \beta_j(v_j x_j(\xi_j - \lambda_j) - \bar{v}_j \bar{x}_j(\bar{\xi}_j - \bar{\lambda}_j)) - \hat{\beta}_j(v_j x_j \eta_j - \bar{v}_j \bar{x}_j \bar{\eta}_j) \\ &= \beta_j(\xi_j \bar{v}_j(x_j - \bar{x}_j) + x_j \xi_j(v_j - \bar{v}_j) + \bar{v}_j \bar{x}_j(\xi_j - \bar{\xi}_j)) \\ &\quad - \beta_j(\lambda_j \bar{v}_j(x_j - \bar{x}_j) + x_j \lambda_j(v_j - \bar{v}_j) + \bar{v}_j \bar{x}_j(\lambda_j - \bar{\lambda}_j)) \\ &\quad - \hat{\beta}_j(\eta_j \bar{v}_j(x_j - \bar{x}_j) + x_j \eta_j(v_j - \bar{v}_j) + \bar{v}_j \bar{x}_j(\eta_j - \bar{\eta}_j)) \end{aligned}$$

and

$$e_3 - \bar{e}_3 = \gamma_1 d_1 \eta_1(y_1 - \bar{y}_1) + \gamma_1 d_1 \bar{y}_1(\eta_1 - \bar{\eta}_1) + \gamma_2 d_2 \eta_2(y_2 - \bar{y}_2) + \gamma_2 d_2 \bar{y}_2(\eta_2 - \bar{\eta}_2).$$

Then

$$\begin{aligned} &\|L(u_1, u_2) - L(\bar{u}_1, \bar{u}_2)\|_{L^\infty(0,A)\times L^\infty(0,A)} \\ &\le \frac{C_4}{2B_1}\Big(\|x_1 - \bar{x_1}\|_{L^\infty(0,A)} + \|x_2 - \bar{x_2}\|_{L^\infty(0,A)} + \|v_1 - \bar{v_1}\|_{L^\infty(0,A)} \\ &\quad + \|v_2 - \bar{v_2}\|_{L^\infty(0,A)}\Big) + \frac{C_4}{2B_1}\Big(\|\xi_1 - \bar{\xi}_1\|_{L^\infty(0,A)} + \|\xi_2 - \bar{\xi}_2\|_{L^\infty(0,A)} \\ &\quad + \|\lambda_1 - \bar{\lambda}_1\|_{L^\infty(0,A)} + \|\lambda_2 - \bar{\lambda}_2\|_{L^\infty(0,A)}\Big) \\ &\quad + \big(\frac{C_4}{2B_1} + \frac{C_5}{2B_2}\big)(\|\eta_1 - \bar{\eta}_1\|_{L^\infty(0,A)} + \|\eta_2 - \bar{\eta}_2\|_{L^\infty(0,A)}) + \frac{C_6}{2B_2}\|y_1 - \bar{y}_1\|_{L^\infty(0,A)} \\ &\quad + \frac{C_6}{2B_2}\|y_2 - \bar{y}_2\|_{L^\infty(0,A)} + \frac{1}{2}\big(\frac{A_2}{B_1} + \frac{A_3}{B_2}\big)(\|i_1 - \bar{i_1}\|_{L^\infty(Q)} + \|i_2 - \bar{i_2}\|_{L^\infty(Q)}). \end{aligned}$$

Using the Lipschitz properties of the state and adjoint systems in Theorems 4.1 and 4.4, respectively, we have

$$\|L(u_1, u_2) - L(\bar{u}_1, \bar{u}_2)\| \le \frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)\big(\|u_1 - \bar{u}_1\|_{L^\infty(0,A)} + \|u_2 - \bar{u}_2\|_{L^\infty(0,A)}\big). \tag{4.30}$$

If $\frac{\bar{C}_{A,T}}{2}(\frac{1}{B_1} + \frac{1}{B_2}) < 1$, then the map L admits a unique fixed point (u_1^*, u_2^*), by the Banach Contraction Theorem. Next, we show that this fixed point is an optimal control pair, by using the minimizers, $(u_1^\varepsilon, u_2^\varepsilon)$, from Ekeland's Principle. To do this, we use the states $(x_1^\varepsilon, x_2^\varepsilon, y_1^\varepsilon, y_2^\varepsilon, V_1^\varepsilon, V_2^\varepsilon, S^\varepsilon, i_1^\varepsilon, i_2^\varepsilon)$ and $(\lambda_1^\varepsilon, \lambda_2^\varepsilon, \xi_1^\varepsilon, \xi_2^\varepsilon, \eta_1^\varepsilon, \eta_2^\varepsilon, p^\varepsilon, q_1^\varepsilon, q_2^\varepsilon)$ corresponding to the minimizer $(u_1^\varepsilon, u_2^\varepsilon)$. Thus

$$\begin{aligned} &\big\|L(u_1^\varepsilon, u_2^\varepsilon) - \mathcal{H}\Big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_1}, \frac{e_3^\varepsilon - A_3K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_2}\Big)\big\|_{(L^\infty(0,A))^2} \\ &= \big\|\mathcal{H}\Big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon}{2B_1}, \frac{e_3^\varepsilon - A_3K^\varepsilon}{2B_2}\Big) \\ &\quad - \mathcal{H}\Big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_1}, \frac{e_3^\varepsilon - A_3K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_2}\Big)\big\|_{(L^\infty(0,A))^2} \\ &\le \|\frac{\sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_1}\|_{L^\infty(0,A)} + \|\frac{\sqrt{\varepsilon}\kappa_2^\varepsilon}{2B_2}\|_{L^\infty(0,A)} = \frac{\sqrt{\varepsilon}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big). \end{aligned} \tag{4.31}$$

Next, we show that $(u_1^\varepsilon, u_2^\varepsilon) \to (u_1^*, u_2^*)$ in $L^\infty(0, A) \times L^\infty(0, A)$. Now,

$$\begin{aligned} &\|(u_1^*, u_2^*) - (u_1^\varepsilon, u_2^\varepsilon)\|_{(L^\infty(0,A))^2} \\ &= \|u_1^* - u_1^\varepsilon\|_{L^\infty(0,A)} + \|u_2^* - u_2^\varepsilon\|_{L^\infty(0,A)} \\ &= \big\|\mathcal{H}_1\big(\frac{a_1^* + a_2^* - A_2K^*}{2B_1}\big) - \mathcal{F}_1\big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon - \sqrt{\varepsilon}\kappa_1^*}{2B_1}\big)\big\|_{L^\infty(0,A)} \end{aligned}$$

$$
\begin{aligned}
&+ \big\|\mathcal{H}_2\big(\frac{a_3^* - A_3K^*}{2B_2}\big) - \mathcal{F}_2\big(\frac{e_3^\varepsilon - A_3K^\varepsilon - \sqrt{\varepsilon}\kappa_2^\varepsilon}{2B_2}\big)\big\|_{L^\infty(0,A)} \\
&= \big\|L(u_1^*, u_2^*) - \mathcal{H}\big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_1}, \frac{e_3^\varepsilon - A_3K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_2}\big)\big\|_{(L^\infty(0,A))^2} \\
&\le \|L(u_1^*, u_2^*) - L(u_1^\varepsilon, u_2^\varepsilon)\|_{L^\infty(0,A)} \\
&\quad + \big\|L(u_1^\varepsilon, u_2^\varepsilon) - \mathcal{H}\Big(\frac{e_1^\varepsilon + e_2^\varepsilon - A_2K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_1}, \frac{e_3^\varepsilon - A_3K^\varepsilon - \sqrt{\varepsilon}\kappa_1^\varepsilon}{2B_2}\Big)\big\|_{L^\infty(0,A)} \\
&\le \frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)\big(\|u_1^* - u_1^\varepsilon\|_{L^\infty(0,A)} + \|u_2^* - u_2^\varepsilon\|_{L^\infty(0,A)}\big) + \frac{\sqrt{\varepsilon}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big),
\end{aligned}
$$

from equations (4.30) and (4.31). Also, a_j^* and e_j^* are defined in equations (4.28) and (4.29), respectively. Thus,

$$
\begin{aligned}
&\|u_1^* - u_1^\varepsilon\|_{L^\infty(0,A)} + \|u_2^* - u_2^\varepsilon\|_{L^\infty(0,A)} \\
&\le \frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)\big(\|u_1^* - u_1^\varepsilon\|_{L^\infty(0,A)} + \|u_2^* - u_2^\varepsilon\|_{L^\infty(0,A)}\big) + \frac{\sqrt{\varepsilon}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big).
\end{aligned}
$$

Whence,

$$
\|u_1^* - u_1^\varepsilon\|_{L^\infty(0,A)} + \|u_2^* - u_2^\varepsilon\|_{L^\infty(0,A)} \le \frac{\frac{\sqrt{\varepsilon}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)}{1 - \frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)},
$$

for $\frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)$ sufficiently small. Equivalently,

$$
\|(u_1^*, u_2^*) - (u_1^\varepsilon, u_2^\varepsilon)\|_{L^\infty(0,A)\times L^\infty(0,A)} \le \frac{\frac{\sqrt{\varepsilon}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)}{1 - \frac{\bar{C}_{A,T}}{2}\big(\frac{1}{B_1} + \frac{1}{B_2}\big)} \to 0 \quad \text{as } \varepsilon \to 0^+.
$$

Thus,

$$
(u_1^\varepsilon, u_2^\varepsilon) \to (u_1^*, u_2^*) \quad \text{in } L^\infty(0,A) \times L^\infty(0,A).
$$

Lastly, we establish that (u_1^*, u_2^*) is indeed a minimizer of the functional, $\mathcal{J}$. Now, using Ekeland's Principle, we have $\mathcal{J}(u_1^\varepsilon, u_2^\varepsilon) \le \inf_{(u_1,u_2)\in\mathcal{U}} \mathcal{J}(u_1, u_2) + \varepsilon$. Since $(u_1^\varepsilon, u_2^\varepsilon) \to (u_1^*, u_2^*)$ as $\varepsilon \to 0^+$, it follows that $\mathcal{J}(u_1^*, u_2^*) \le \inf_{(u_1,u_2)\in\mathcal{U}} \mathcal{J}(u_1, u_2)$. □

5. Numerical simulations

We present a numerical scheme for the within-host model (2.1)–(2.3) and the between-host model (2.4)–(2.9) based on semi-implicit finite difference schemes for ordinary and first-order partial differential equations [2]. Let $\Delta\tau = h > 0$ be the discretization step for the interval $[0, A]$, with $h = A/M$, where M is the total number of subintervals in age (age-since-infection), and $\Delta t = k > 0$ be the discretization step for the interval $[0, T]$, with $k = \frac{T}{N}$, where N is the total number of subintervals in time. We discretize the intervals $[0, A]$ and $[0, T]$ at the points $\tau_j = j\Delta\tau$ $(j = 0, 1, \ldots, M)$ and $t_n = n\Delta t$ $(n = 0, 1, \ldots, N)$, respectively. Next, we define the state, adjoint and control functions in terms of nodal points x_1^j, x_2^j, y_1^j, y_2^j, v_1^j, v_2^j, S^n, ω_j^n (where $\omega \equiv i_1$), $\tilde{\omega}_j^n$ (where $\tilde{\omega} \equiv i_2$), ψ_1^j, ψ_2^j, φ_1^j, φ_2^j, ϕ_1^j, ϕ_2^j, θ^n, λ_j^n, $\tilde{\lambda}_j^n$, u_1^j and u_2^j. Since ω_j^n is an approximation to the solution of the equation

that models infectious individuals of group one at time level t_n and grid point τ_j, we approximate the directional derivatives $\frac{\partial\omega(\tau,t)}{\partial t}$ and $\frac{\partial\omega(\tau,t)}{\partial\tau}$ by

$$\frac{\partial\omega(\tau_j,t_n)}{\partial t}\approx\frac{\omega_j^n-\omega_j^{n-1}}{\Delta t},\quad \frac{\partial\omega(\tau_j,t_n)}{\partial\tau}\approx\frac{\omega_j^{n-1}-\omega_{j-1}^{n-1}}{\Delta\tau}.$$

Age of individuals changes at the same speed as chronological time, and therefore we assume that $\Delta t=\Delta\tau$, so that

$$\frac{\partial\omega(\tau_j,t_n)}{\partial t}+\frac{\partial\omega(\tau_j,t_n)}{\partial\tau}\approx\frac{\omega_j^n-\omega_{j-1}^{n-1}}{\Delta t}.$$

We fully implement our numerical scheme for the multi-group coupled within-host and between-host model by using parameter values of the within-host and epidemiological model of HIV given in Table 1. For this set of parameter values, the basic reproduction number of the epidemiological model in the absence of control is $\mathcal{R}_0=3.9$, and in the presence of drug treatment is $\tilde{\mathcal{R}}_0=2.1$. Here, $\tilde{\mathcal{R}}_0$ denotes the basic reproduction of the epidemiological model in the presence of drug treatment on the within-host system.

TABLE 1. Within-host and between-host parameter values

Parameter	Value	Units	Source
r	10	cells mm^{-3}day^{-1}	[16, 27, 35]
m_0	0.012	mm^3 year^{-1}	[34]
Λ	2750	humans	[34]
μ	0.02	day^{-1}	[35, 39]
p_1	0.7	–	vary
p_2	0.3	–	vary
β_1	2.4 $\times10^{-5}$	mm^{3}virion^{-1}day^{-1}	[27, 35, 39]
$\hat{\beta}_1$	2.4 $\times10^{-5}$	mm^{3}cell^{-1}day^{-1}	[27, 35, 39]
d_1	0.5	day^{-1}	[27, 35]
ν_1	1200	virions cell^{-1}	[16]
δ_1	2.5	day^{-1}	[16, 35]
s_1	0.014	day^{-1}	[34]
c_1	4×10^{-5}	mm^{3}virion^{-1}year^{-1}	[34]
μ_1	2×10^{-7}	virion^{-1}year^{-1}	[34]
β_2	2.0 $\times10^{-5}$	mm^{3}virion^{-1}day^{-1}	[16, 35, 39]
$\hat{\beta}_2$	2.0 $\times10^{-5}$	mm^{3}cell^{-1}day^{-1}	[16, 35, 39]
d_2	0.5	day^{-1}	[16, 35]
ν_2	1200	virions cell^{-1}	[16]
δ_2	3	day^{-1}	[16, 35]
s_2	1.4	day^{-1}	[34]
c_2	4×10^{-5}	mm^{3}virion^{-1}year^{-1}	[34]
μ_2	2×10^{-7}	virion^{-1}year^{-1}	[34]

Starting with 600 healthy cells for both groups of healthy cells at the within-host level, no infectious cells ($y_1(0)=y_2(0)=0$), but with different viral loads, Figure 1 delineates trajectories for the within-host dynamics within a time horizon of 100 days. With a "higher" viral load of $v_1(0)=0.005$ for free virus of group one and a

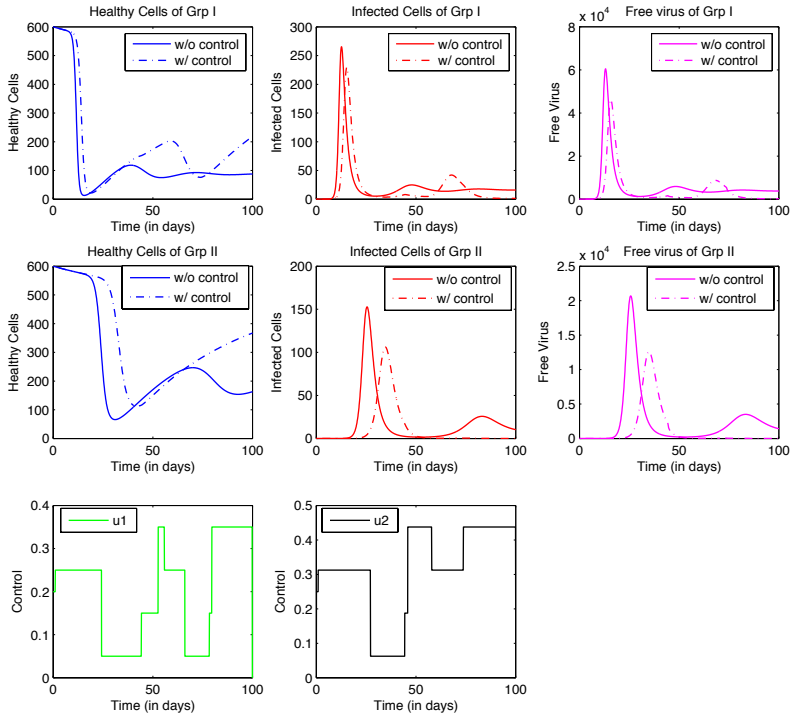

FIGURE 1. Within-host Dynamics when $x_1(0) = 600$, $x_2(0) = 600$, $y_1(0) = y_2(0) = 0$, $v_1(0) = 0.005$, $v_2(0) = 0.001$, $A_1 = 1$, $A_2 = 1$, $A_3 = 1$, $A_4 = 1$, $B_1 = 1$, $B_2 = 1$, $\tilde{u}_1 = 0.4$ and $\tilde{u}_2 = 0.5$.

"lower" viral load of $v_2(0) = 0.001$ for the free virus of group two, trajectories for healthy cells of group one indicate a rapid decrease within the first twenty days for group one, and a decrease within the first thirty days for healthy cells of group two. For the free virus population of both groups, acute phases are observed in different groups, but within different time horizons. Free virus of group one observes an acute phase between 10–20 days since start-of-infection as opposed to 20–40 days since start-of-infection for free virus of group two. Also, relapse phases are observed in both groups of the free virus. For free virus of group one, the relapse phase occurs within 50 days since start-of-infection and within 90 days since start-of-infection for free virus of group two. In the presence of fusion and protease inhibitors, the acute and relapse phases of the virus of group one occurs much later. However, the acute phase for free virus of group two occurs much later, but with no relapse phase within 100 days since start-of-infection.

At the population level, and starting with initial age distributions of $i_1(\tau, 0) = 200 \sin(\frac{\pi\tau}{25})$ and $i_2(\tau, 0) = 200 \sin(\frac{\pi\tau}{25})$ for infectious individuals of both groups, and an initial population of $S(0) = 1 \times 10^6$ for susceptible individuals, oscillatory behaviors are observed in both populations as shown in Figure 2. Due to higher viral load for free virus of group one, more infectious cases are also observed at

the population level for infectious individuals of group one in the absence of drug treatment on the within-host system. In the presence of drug treatment, there is an oscillatory increase and decrease in the number of infectious cases, but with more infectious cases observed in the presence of control as compared to the the number obtained in the absence of control in both groups. This may be attributed to the fact that, infectious individuals tend to live longer in the presence of drugs than in the absence of drugs.

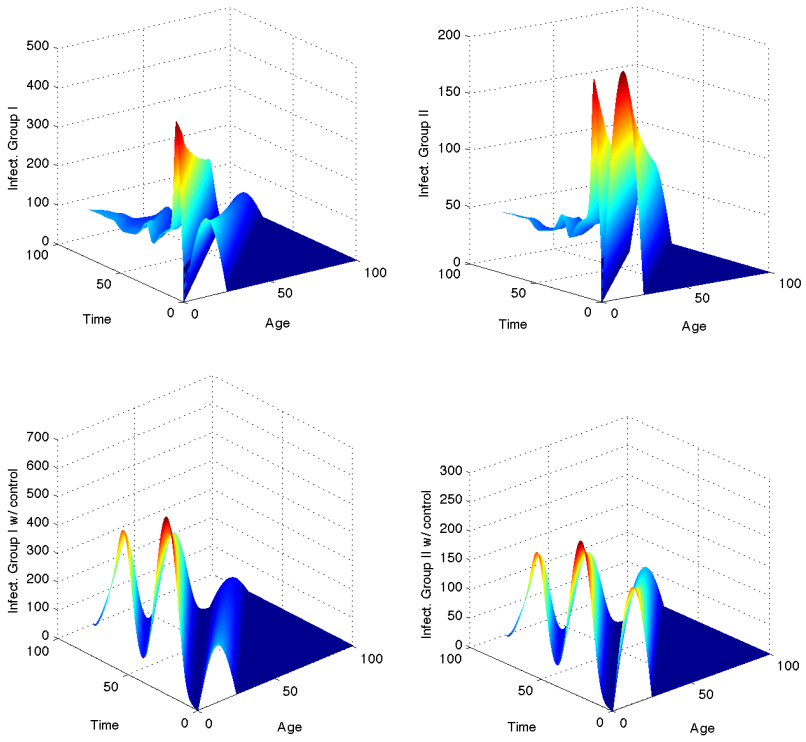

FIGURE 2. Infectious Individuals when $i_1(\tau,0) = 200\sin(\frac{\pi\tau}{25})$, $i_2(\tau,0) = 200\sin(\frac{\pi\tau}{25})$, $S(0) = 1\times 10^6$, $A_1 = 1$, $A_2 = 1$, $A_3 = 1$, $A_4 = 1$, $B_1 = 1$, $B_2 = 1$, $\tilde{u}_1 = 0.4$ and $\tilde{u}_2 = 0.5$.

In Figure 3, trajectories depict susceptible individuals in the absence and presence of drug treatment on the within-host system. In the absence of control, susceptible individuals experience a decrease in population over the entire time horizon. In the presence of drug treatment, susceptible individuals still experience a decrease in population, but with more susceptible cases observed in the population.

Figure 4 represents trajectories for the within-host dynamics when the effectiveness of the fusion and protease inhibitors is very high. With this level of effectiveness, the number of healthy cells of group one experiences an increase between

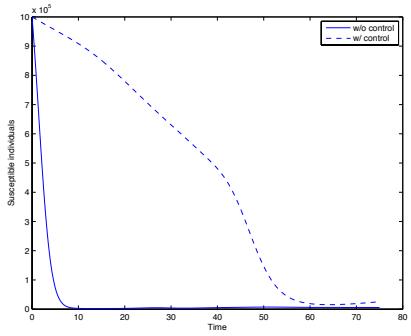

FIGURE 3. Susceptible Individuals when $i_1(\tau,0) = 200\sin(\frac{\pi\tau}{25})$, $i_2(\tau,0) = 200\sin(\frac{\pi\tau}{25})$ and $S(0) = 1\times 10^6$.

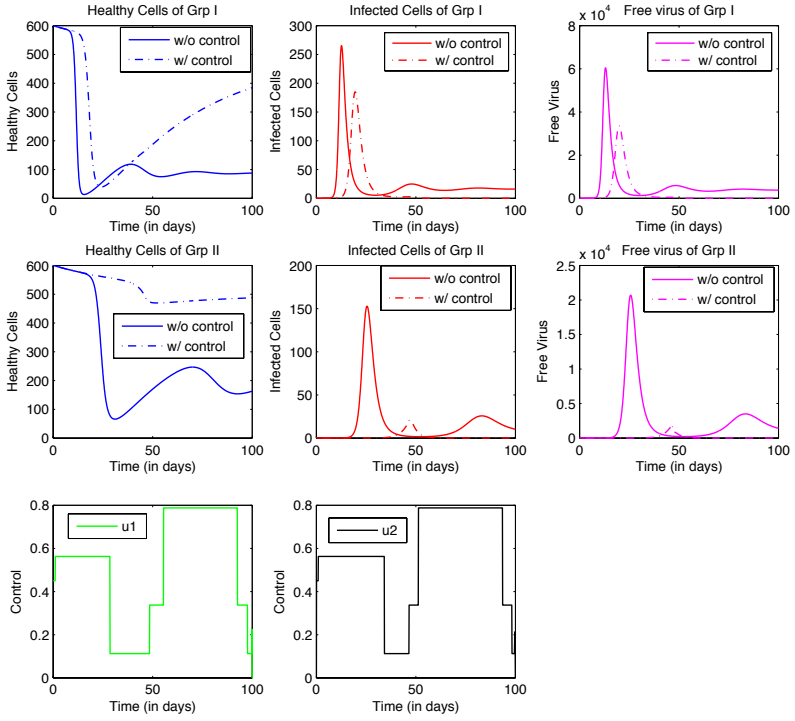

FIGURE 4. Within-host Dynamics when $x_1(0) = 600$, $x_2(0) = 600$, $y_1(0) = y_2(0) = 0$, $v_1(0) = 0.005$, $v_2(0) = 0.001$, $\tilde{u}_1 = 0.9$ and $\tilde{u}_2 = 0.9$.

25–100 days since start-of-infection, and healthy cells of group two experiences a subtle decrease followed by an increase in the number of healthy cells within the

rest of the time horizon. For the population of free virus of both groups, the relapse phase observed in the absence of control is not observed in the presence of control. The control suggests an intermediate level of treatment within the first 30 days since start-of-infection, followed by a high level of treatment between 50–95 days since start-of-infection. At the population level, and considering the total population of infectious individual of both groups, numerical simulations suggest that the disease could be controlled as indicated in Figure 5.

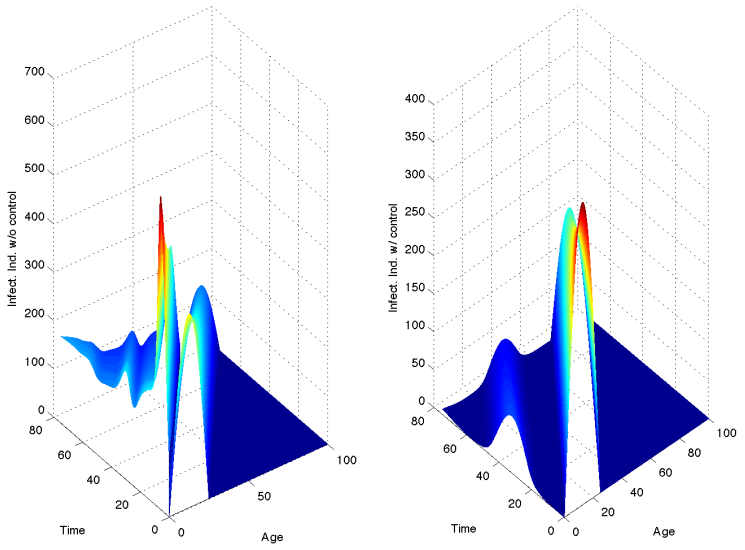

FIGURE 5. Total population of infectious individuals when $i_1(\tau,0) = 200\sin(\frac{\pi\tau}{25})$, $i_2(\tau,0) = 200\sin(\frac{\pi\tau}{25})$ and $S(0) = 1\times 10^6$.

Conclusion. In this study, we have formulated in a careful mathematical way, a multi-group within-host model coupled with an epidemiolgical model. Explicit dependence of the epidemiological model on within-host dynamics are expressed in transmission and mortality rates at the population level. Existence of solution is established via a fixed point argument. The basic reproduction number of the multi-group epidemiological model is derived and an explicit dependence on the within-host viral load is captured. Global stability analysis of disease-free equilibrium and local asymptotic stability analysis of endemic equilibrium are obtained.

We formulated an optimal control problem for the coupled model subject to fusion and protease inhibitors. Sensitivity and adjoint systems are derived, and existence, characterization and uniqueness results obtained. Using a semi-implicit finite difference scheme on the state and adjoint systems, and a forward-backward sweep iterative method, the optimality system is solved numerically. Numerical simulations suggest that the combination of fusion and protease inhibitors reduces viral load at the within-host level and the disease-induced mortality at the population level, but results in an increase in the number of infectious individuals at

the population level since infectious individuals live longer in the presence of drugs. The disease could still be controlled if the effectiveness of treatment is at a very high level. At this level of control, the basic reproduction number of the epidemiological model in the presence of drug treatment reduces to $\tilde{\mathcal{R}}_0 = 0.19$.

Acknowledgments. S. Lenhart would like to acknowledge partial support from the National Institute for Mathematical and Biological Synthesis through the National Science Foundation Award # EF–0832858, and partial support from the Center for Business and Economic Research at the University of Tennessee. The work of Numfor was supported in part by the College of Science and Mathematics at Georgia Regents University. Maia Martcheva acknowledges partial support from grant DMS-1515661.

References

[1] R. M. Anderson, R. M. May; The Invasion, Persistence and Spread of Infectious Diseases Within Animal and Plant Communities, *Phil. Trans. R. Soc. Lond.*, 314 (1986), 533–570.

[2] S. Anita, Viorel Arnăutu, Vincenzo Capasso; *An Introduction to Optimal Control Problems in Life Sciences*, Springer Science, New York, 2011.

[3] S. Anita; *Analysis and Control of Age-Dependent Population Dynamics*, Kluwer Academic Publishers, Dordretcht, 2000.

[4] S. Anita; Optimal Control Harvesting for a Nonlinear Age-dependent Population Dynamics, *J. Math. Anal. Appl.* 226 (1998), 6–22.

[5] R. Antia, B. Levin, R. M. May; Within-host Population Dynamics and the Evolution and Maintenance of Microparasite Virulence, *Am. Nat.*, 144 (1994), 457–472.

[6] V. Barbu, M. Iannelli; Optimal Control of Population Dynamics, *J. Optim. Theory Appl.*, 102 (1999),1–14.

[7] V. Barbu; *Mathematical Methods in Optimization of Differential Systems*, Kluwer Academic Publishers, Dordretcht, 1994.

[8] M. Brokate; Pontryagin's Principle for Control Problems in Age-Dependent Population Dynamics, *J. Math. Biol.*, 23 (1985), 75–101.

[9] C. Castillo-Chavez, Z. Feng; Global Stability of an Age-Structured Model for TB and its Applications to Optimal Vaccination Strategies, *Math.Biosci.*, 151 (1998), 135–154.

[10] C. Castillo-Chavez, H. W. Hethcote, V. Andreasen, S. A. Levein, W. M. Liu; Epidemiological Models with Age Structure, Proportionate Mixing, and Cross-immunity, *J. Math. Biol.* 27 (1989), 233–258.

[11] D. Coombs, M. A. Gilchrist, J. Percus, A. S. Perelson; Optimizing Viral Production, *Bull. Math. Biol.*, 65 (2003) 1003–1023.

[12] U. Dissanayake, S. Dias, H. Polson, S. Longacre, P. Udagama-Randeniya; Immuno-Epidemiology of Plasmodium Vivax Merozoite Surface Proten-4, *Bio. Sci.*, 37 (2008) 97–105.

[13] I. Ekeland; On the Variational Principle, *J. Math. Appl.*, 47 (1974),324–353.

[14] Z. Feng, J. Velasco-Hernandez, B. Tapia-Santos; A Mathematical model for Coupling within-Host and Between-Host Dynamics in an Environmentally-driven Infectious Disease, *Math. Biosci.*, 241 (2013), 49–55.

[15] Z. Feng, J. Velasco-Hernandez, B. Tapia-Santos, M. C. Leite; A Model for Coupled Within-Host and Between-Host Dynamics in an Infectious Disease, *Nonlinear Dyn.*, 68 (2012), 401–411.

[16] K. R. Fister, S. Lenhart, J. S. McNally; Optimizing Chemotherapy in an HIV Model, *J. of Differential Equations*, 1998 (1998), 1–12.

[17] K. R. Fister, S. Lenhart; Optimal Harvesting in an Age-Structured Predator-Prey Model, *Appl. Math. Optim.*, 54 (2006), 1–15.

[18] K. R. Fister, S. Lenhart; Optimal Control of a Competitive System with Age-Structured, *J. Math. Anal.*, 291 (2004), 526–537.

[19] V. V. Ganusov, C. T. Bergstrom, R. Antia; Within-Host Population Dynamics and the Evolution of Microparasites in a Heterogeneous Host Population, *Evolution*, 56 (2002), no. 2, 213–223.

[20] M. A. Gilchrist, D. A. Coombs; Evolution of Virulence: Interdependence, Constraints, and Selection using Nested models, *Theor. Pop. Biol.*, 69 (2006), 145–153.
[21] M. A. Gilchrist, A. Sasaki; Modeling Host-Parasite Coevolution: A Nested Approach based on Mechanistic Models, *J. Theor. Biol.*, 218 (2002), 289–308.
[22] M. E. Gurtin, R. C. MacCamy; Nonlinear Age-dependent Population Dynamics, *Arch. Rat. Mech. Anal.*, 54 (1974), 281–300.
[23] J. M. Heffernan; Mathematical Immunology of Infectious Diseases, *Math. Popul. Stud.*, 18 (2011), 47–54.
[24] B. Hellriegel; Immunoepidemiology-Bridging the Gap Between Immunology and Epidemiology, *Trends in Parasitology*, 17 (2001), 102–106.
[25] M. Iannelli, M. Martcheva, F. A. Milner; *Gender-Structured Population Modeling: Mathematical Methods, Numerics, and Simulations*, SIAM, Philadelphia, 2005.
[26] H. Joshi; Optimal Control of an HIV Immunology Model, *Optimal Control Appl. Methods*, 23 (2002), 199–213.
[27] D. Kirschner, S. Lenhart, S. Serbin; Optimal Control of the Chemotherapy of HIV, *J. Math. Biol.*, 35 (1997), 775–792.
[28] S. Lenhart, J. T. Wortman; *Optimal Control Applied to Biological Models*, Taylor & Francis, Boca Raton, FL, 2007.
[29] X. Li, J. Liu, M. Martcheva; An Age-Structured Two-strain Epidemic Model with Superinfection, *Math. Biosci. Eng.*, 7 (2010), 123–147.
[30] M. Martcheva; An Immuno-epidemiological Model of Paratuberculosis, *AIP Conf. Proc.*, 1404 (2011), 176–183.
[31] M. Martcheva and X. Li, Linking Immunological and Epidemiological Dynamics of HIV: the case of super-infection, *J. Biol. Dyn.*, 7 (2013), 161–182.
[32] M. Martcheva, F. A. Milner; A Two-Sex Age-Structured Population Model: Well Posedness, *Math. Population Studies*, 7 (1999), 111–129.
[33] L. F. Murphy, S. J. Smith; Optimal Harvesting of an Age-Structured Population, J. Math. Biol., 29 (1990), 77–90 .
[34] E. Numfor, S. Bhattacharya, S. Lenhart, M. Martcheva; Optimal Control in Coupled Within-host and Between-host Models, *Math. Model. Nat. Phenom.*, 9 (2014), 171–203.
[35] A. Perelson, D. E. Kirschner, R. D. Boer; Dynamics of HIV Infection of $CD4^+$ T Cells, *Math. Biosci.*, 114 (1993), 81–125.
[36] Z. Qiu, X. Li, M. Martcheva; Multi-strain Persistence Induced by Host Age Structure, *J. Math. Anal. Appl.*, 391 (2012), 395–612.
[37] E. Shim, Z. Feng, M. Martcheva, C. Castillo-Chavez; An Age-Structured Epidemic Model of Rotavirus with Vaccination, *Math. Biol.*, 53 (2006), 719–746.
[38] G. F. Webb; *Theory of Nonlinear Age-Dependent Population Dynamics*, Marcel Dekker, Inc, New York, 1985.
[39] E. K. Yeargers, R. W. Shonkwiler, J. V. Herod; *An Introduction to the Mathematics of Biology with Computer Algebra Models*, Birkhäuser, Boston, 1996.

Eric Numfor
Department of Mathematics, Georgia Regents University, Augusta, GA 30912, USA
E-mail address: enumfor@gru.edu

Souvik Bhattacharya
Department of Mathematics, University of Trento, Italy
E-mail address: souvikb2005@gmail.com

Maia Martcheva
Department of Mathematics, University of Florida, Gainesville, FL 32611, USA
E-mail address: maia@ufl.edu

Suzanne Lenhart
Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA
E-mail address: lenhart@math.utk.edu