OPTIMAL INTERVENTION STRATEGIES OF STAGED PROGRESSION HIV INFECTIONS THROUGH AN AGE-STRUCTURED MODEL WITH PROBABILITIES OF ART DROP OUT

MBOYA BA[1], RAMSÈS DJIDJOU-DEMASSE[2,*], MOUNTAGA LAM[1]
AND JEAN-JULES TEWA[3]

Abstract. In this paper, we construct a model to describe the transmission of HIV in a homogeneous host population. By considering the specific mechanism of HIV, we derive a model structured in three successive stages: (i) primary infection, (ii) long phase of latency without symptoms, and (iii) AIDS. Each HIV stage is stratified by the duration for which individuals have been in the stage, leading to a continuous age-structure model. In the first part of the paper, we provide a global analysis of the model depending upon the basic reproduction number $\mathcal{R}_0$. When $\mathcal{R}_0 \leq 1$, then the disease-free equilibrium is globally asymptotically stable and the infection is cleared in the host population. On the contrary, if $\mathcal{R}_0 > 1$, we prove the epidemic's persistence with the asymptotic stability of the endemic equilibrium. By performing the sensitivity analysis, we then determine the impact of control-related parameters on the outbreak severity. For the second part, the initial model is extended with intervention methods. By taking into account antiretroviral therapy (ART) interventions and the probability of treatment drop out, we discuss optimal intervention methods which minimize the number of AIDS cases.

Mathematics Subject Classification. 35Q92, 49J20, 35B35, 92D30.

1. INTRODUCTION

1.1. Biology and evolution of HIV infection

The human immunodeficiency virus (HIV) is a virus that attacks the immune system, the body's defense against infection. HIV weakens your immune system by destroying cells that are essential for fighting diseases and infections. Without treatment, the immune system becomes too weak. A chronic progressive disease called AIDS (Acquired Immunodeficiency Syndrome) then appears. The situation of the epidemic in the world shows only a stabilization of the number of new cases diagnosed, although undeniable efforts have been made in recent years [49]. The complexity of HIV infection is linked to many elements that particularly involve the specific mechanism of infection [22]. In the absence of treatment, the HIV infection goes through three successive stages

[1] University Cheikh Anta Diop, Department of Mathematics and Informatics, Faculty of Science and Technic, Dakar, Sénégal.
[2] MIVEGEC, Univ. Montpellier, IRD, CNRS, Montpellier, France.
[3] University of Yaounde I, National Advanced School of Engineering, Yaoundé, Cameroon.

* Corresponding author: ramses.djidjoudemasse@ird.fr

corresponding to T4 cell count ranges: (i) primary infection, (ii) long phase of latency without symptoms, and (iii) AIDS [52]. *Primary infection:* The risk of transmission is particularly high during this phase because of the high viral load at this stage of the infection [21, 40, 51]. This stage is characterized by occurrence of symptoms similar to those of a cold or mild influenza (fever, rashes, fatigue, headaches) which disappear spontaneously after a few weeks [1, 51]. *Phase of latency without symptoms:* Generally, HIV-positive people do not experience any particular problem at this stage for many years and can lead normal lives, although the virus is spreading insidiously in the body and permanently mistreats the immune system [1, 51]. *AIDS stage:* Because of its constant solicitation, the immune system becomes weaker and weaker until it can no longer defend itself against many pathogen agents and prevent the occurrence of serious or fatal diseases [1, 51].

1.2. Waiting time within HIV infection stages

Staged progression models have been proposed to investigate the transmission dynamics of HIV [13, 18, 19, 23, 24, 30, 35, 39]. Here, we go through the same direction by modeling the progression through the three HIV stages described previously (which are enough for the practical interpretation of the HIV stages up to date). However, none of the aforementioned works deal with a continuous stage-structured model as in the context of this work. Indeed, here we consider the duration $a \geq 0$ spent in a given HIV stage as a continuous variable (not to be confused with the time since infection as in [44], or with age group such as "youths", "adults", etc). Further, our model formulation is well adapted for the dynamics of HIV infection without any treatment: the reference values for the duration of HIV stages 1 and 2 are 2.90 (range 1.23–6) and 120 (range 108–180) months respectively [21]. Moreover, the model proposed here is suitable for the specific mechanisms of antiretroviral therapy, ART for short, which is the use of HIV medicines to slow down the progress of the infection. ART can help people with HIV live longer by extending the time spent in a given HIV stage [2].

1.3. The model

Here we formulate a model structured by the duration a for which individuals have been in a given HIV stage. The model is called *age-structured* model thereafter. The host population is divided into four compartments: $S(t)$ denotes susceptible individuals at time t, while $i_1(t, a)$, $i_2(t, a)$ and $i_3(t, a)$ respectively denote the number of the primary infected, infected in asymptomatic stage and infected in AIDS stage at time t who have been in that stage for duration $a > 0$ (the time spent in the stage). The transfer diagram reads:
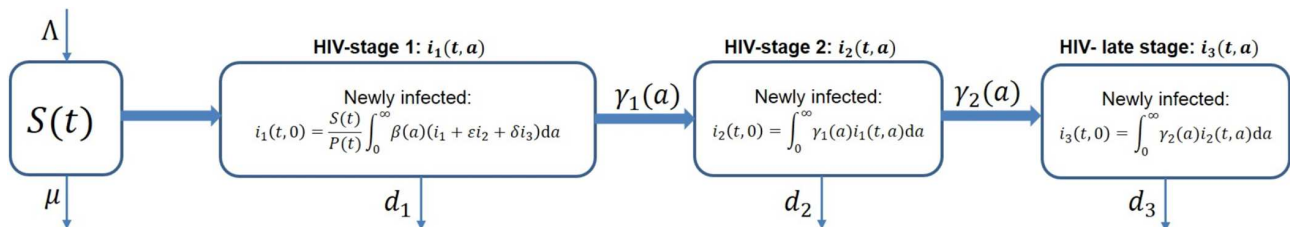
The parameter Λ represents the positive contribution entering into the susceptible population and μ is the natural death rate of susceptible individuals. The function $\beta = \beta(a)$ is the rate of being infectious after time a within HIV stage 1, $\varepsilon > 0$ and $\delta \geq 0$ are the reduced transmission rate in HIV stages 2 and 3. Note that the reduced transmission rates, ε and δ, can be considered duration-dependent ($\varepsilon = \varepsilon(a)$ and $\delta = \delta(a)$) with no more difficulties in the analysis proposed here. Our choice of constant reduced rate is mostly motivated by the biological knowledge of HIV-stages transmission efficiencies [13, 21]. The death rate of infectious individuals i_1, i_2 and i_3, are respectively denoted by functions $d_1 = d_1(a)$, $d_2 = d_2(a)$ and $d_3 = d_3(a)$. Obviously, infectious individuals are assumed to have an increased rate of death (*i.e.* $d_j \geq \mu$). The rate of disease progression from infectious class i_1 to the infectious class i_2 is $\gamma_1 = \gamma_1(a)$, as well as $\gamma_2 = \gamma_2(a)$ the rate of progression from infectious class i_2 to the infectious class i_3.

The model we shall consider in this work reads as the following age-structured system of equations,

$$
\begin{cases}
\dot{S}(t) = \Lambda - \mu S(t) - \frac{S(t)}{P(t)} \int_0^\infty \beta(a) \left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a)\right) \mathrm{d}a, \\
i_1(t,0) = \frac{S(t)}{P(t)} \int_0^\infty \beta(a) \left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a)\right) \mathrm{d}a, \\
(\partial_t + \partial_a) \, i_1(t,a) = - \left(\gamma_1(a) + d_1(a)\right) i_1(t,a), \\
i_2(t,0) = \int_0^\infty \gamma_1(a) i_1(t,a) \mathrm{d}a, \\
(\partial_t + \partial_a) \, i_2(t,a) = - \left(\gamma_2(a) + d_2(a)\right) i_2(t,a), \\
i_3(t,0) = \int_0^\infty \gamma_2(a) i_2(t,a) \mathrm{d}a, \\
(\partial_t + \partial_a) \, i_3(t,a) = -d_3(a) i_3(t,a), \\
\text{with} \\
P(t) = S(t) + \int_0^\infty \left(i_1(t,a) + i_2(t,a) + i_3(t,a)\right) \mathrm{d}a,
\end{cases}
\tag{1.1}
$$

coupled with the initial condition

$$
S(0) = S_0, i_1(0,a) = i_{10}(a), i_2(0,a) = i_{20}(a), i_3(0,a) = i_{30}(a).
\tag{1.2}
$$

Furthermore, model parameters are assumed to satisfy the following hypotheses:

Assumption 1.1.   1. $\Lambda, \mu > 0$; $\beta, d_1, d_2, d_3, \gamma_1, \gamma_2, \in L_+^\infty(0,\infty)$ and $d_j(\cdot) \geq \mu$.
   2. β and γ_j are Lipschitz continuous almost everywhere on $\mathbb{R}_+$.
   3. For any $a > 0$, there exists $a_\beta; a_{\gamma_j} > a$, with $j = 1, 2, 3$; such that β is positive in a neighbourhood of a_β and γ_j is positive in a neighbourhood of a_{γ_j}.

1.4. Aims

In the first part of this article, we shall study some dynamical properties of problem (1.1)–(1.2). We shall first compute the basic reproduction number which determines the outcome of the disease and study the stability of the model equilibria. We shall perform the global sensitivity analysis of the HIV late stage. That is to help us know parameters that are most influential in determining disease dynamics, *i.e.* AIDS cases.

In the second part, we then introduce intervention strategies into model (1.1) aiming to optimally reduce AIDS cases in the host population. Several HIV intervention options (called controls) do exist. Individuals do not take the same combination of medicines because the infection stages differ [2]. Therefore, interventions strategies introduced in the model of this note are stage specific. Furthermore, there have been numerous works on the optimal control of age-structured populations [3, 4, 8–10, 14, 16, 17, 29, 36] and references cited therein.

This paper is organized as follows. Section 2 describes the preliminaries results of model (1.1): existence of semiflow and asymptotic behaviour. We derive the global sensitivity analysis, describe the model parameters, and the typical model simulation in Section 3. In Section 4, we extend model (1.1) with intervention strategies and characterize the necessary optimality condition. We then discuss the effectiveness of those interventions and some model hypotheses and limitations in Section 5. Appendices A–D are devoted to the proofs of our main results.

2. PRELIMINARIES

The aim of this section is to provide some preliminary remarks to system (1.1). Let us introduce the following notations, for $a \geq 0$,

$$D_3(a) = \exp\left(-\int_0^a (d_3(\sigma))\,\mathrm{d}\sigma\right), D_j(a) = \exp\left(-\int_0^a (\gamma_j(\sigma) + d_j(\sigma))\,\mathrm{d}\sigma\right); j = 1, 2,$$

$$\overline{D}_j = \int_0^\infty D_j(a)\mathrm{d}a, \quad \Omega_j = \int_0^\infty \beta(\sigma)D_j(\sigma)\mathrm{d}\sigma; \quad j = 1, 2, 3, \tag{2.1}$$

$$\Gamma_j = \int_0^\infty \gamma_j(\sigma)D_j(\sigma)\mathrm{d}\sigma; \quad j = 1, 2.$$

Note that parameters $\overline{D}_j$, Ω_j, and Γ_j,s are finite constants because Assumption 1.1 (item 1) holds. Furthermore, $\overline{D}_j$,s quantify the average lifespan of an individual at the HIV stage-j, Ω_j,s the average infectiousness of an individual at the HIV stage-j, and Γ_j,s the average progression to the HIV stage-$j+1$.

2.1. Existence of semiflow

We first formulate system (1.1) composed by (S, i_1, i_2, i_3) in an abstract Cauchy problem. For that aim, we introduce the Banach space $\mathcal{X} = \mathbb{R} \times \mathbb{R}^3 \times L^1(0, \infty, \mathbb{R}^3)$ endowed with the usual product norm $\|\cdot\|_{\mathcal{X}}$ as well as its positive cone $\mathcal{X}_+$. Let $A : D(A) \subset \mathcal{X} \to \mathcal{X}$ be the linear operator defined by $D(A) = \mathbb{R} \times \{0_{\mathbb{R}^3}\} \times W^{1,1}(0, \infty, \mathbb{R}^3)$ and

$$A\left(v, 0, 0, 0, u_1, u_2, u_3\right)^T = \begin{pmatrix} -\mu v \\ -u_1(0) \\ -u_2(0) \\ -u_3(0) \\ -u_1' - (\gamma_1 + d_1)u_1 \\ -u_2' - (\gamma_2 + d_2)u_2 \\ -u_3' - d_3 u_3 \end{pmatrix}.$$

Finally, let us introduce the non-linear map $F : \overline{D(A)} \to \mathcal{X}$ defined by

$$F\left(v, 0, 0, 0, u_1, u_2, u_3\right)^T = \begin{pmatrix} \Lambda - \frac{v}{v + \sum_{j=1}^3 \int_0^\infty u_j(a)\mathrm{d}a} \int_0^\infty \beta(a)\left(u_1(a) + \varepsilon u_2(a) + \delta u_3(a)\right)\mathrm{d}a \\ \frac{v}{v + \sum_{j=1}^3 \int_0^\infty u_j(a)\mathrm{d}a} \int_0^\infty \beta(a)\left(u_1(a) + \varepsilon u_2(a) + \delta u_3(a)\right)\mathrm{d}a \\ \int_0^\infty \gamma_1(a)u_1(a)\mathrm{d}a \\ \int_0^\infty \gamma_2(a)u_2(a)\mathrm{d}a \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

By identifying $\varphi(t)$ together with $(S(t), 0_{\mathbb{R}^3}, i_1(t, .), i_2(t, .), i_3(t, .))^T$ and by setting $\varphi_0 = (S_0, 0_{\mathbb{R}^3}, i_{10}(.), i_{20}(.), i_{30}(.))^T$, system (1.1) rewrites as the following Cauchy problem

$$\begin{cases} \dfrac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = A\varphi(t) + F(\varphi(t)), \\ \varphi(0) = \varphi_0. \end{cases} \tag{2.2}$$

By setting $\mathcal{X}_0 = \overline{D(A)}$ and $\mathcal{X}_{0+} = \mathcal{X}_0 \cap \mathcal{X}_+$ the precise result is the following theorem.

Theorem 2.1. *Let Assumption 1.1 be satisfied. Then there exists a unique strongly continuous semiflow $\{\Phi(t, \cdot) : \mathcal{X}_0 \to \mathcal{X}_0\}_{t \geq 0}$ such that for each $\varphi_0 \in \mathcal{X}_{0+}$, the map $\varphi \in \mathcal{C}((0, \infty), \mathcal{X}_{0+})$ defined by $\varphi = \Phi(\cdot, \varphi_0)$ is*

a mild solution of (2.2), *namely, it satisfies* $\int_0^t \varphi(s)\mathrm{d}s \in D(A)$ *and* $\varphi(t) = \varphi_0 + A\int_0^t \varphi(s)\mathrm{d}s + \int_0^t F(\varphi(s))\,\mathrm{d}s$ *for all* $t \geq 0$. *Moreover,* $\{\Phi(t,\cdot)\}_t$ *satisfies the following properties:*

1. *Let* $\Phi(t,\varphi_0) = (S(t), 0_{\mathbb{R}^3}, i_1(t,\cdot), i_2(t,\cdot), i_3(t,\cdot))^T$; *then the following Volterra formulation holds true*

$$i_1(t,a) = \begin{cases} i_{10}(a-t)\frac{D_1(a)}{D_1(a-t)}, & \text{for } t < a, \\ \frac{S(t-a)}{P(t-a)}E_1(t-a)D_1(a), & \text{for } t \geq a \end{cases} \tag{2.3}$$

$$i_j(t,a) = \begin{cases} i_{j0}(a-t)\frac{D_j(a)}{D_j(a-t)}, & \text{for } t < a, \\ E_j(t-a)D_j(a), & \text{for } t \geq a, \end{cases} \quad j = 2, 3;$$

wherein

$$E_1(t) = \int_0^\infty \beta(a)\left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a)\right)\mathrm{d}a, \text{ and}$$

$$E_{j+1}(t) = \int_0^\infty \gamma_j(a)i_j(t,a)\mathrm{d}a, \quad j = 1, 2. \tag{2.4}$$

2. *For all* $\varphi_0 \in \mathcal{X}_{0+}$ *one has for all* $t \geq 0$,

$$\|\Phi(t,\varphi_0)\|_{\mathcal{X}} \leq \max\left\{\frac{\Lambda}{\mu}, \frac{\Lambda}{\mu} + e^{-\mu t}\left(\|\varphi_0\|_{\mathcal{X}} - \frac{\Lambda}{\mu}\right)\right\} \leq \max\left\{\frac{\Lambda}{\mu}, \|\varphi_0\|_{\mathcal{X}}\right\}.$$

3. *The semiflow* $\{\Phi(t,\cdot)\}_t$ *is bounded dissipative and asymptotically smooth, that is*

Bounded dissipative: there exists a bounded set $\mathcal{B} \subset \mathcal{X}_0$ *such that for any bounded set* $U \subset \mathcal{X}_0$, *there exists* $\tau = \tau(U, \mathcal{B}) \geq 0$ *such that* $\Phi(t, U) \subset \mathcal{B}$ *for* $t \geq \tau$.

Asymptotically smooth: for any nonempty, closed, bounded set $U \subset \mathcal{X}_0$, *there exists a nonempty compact set* $J = J(U)$ *such that* J *attracts* $\{\varphi \in U : \Phi(t,\varphi) \in U, \forall t \geq 0\}$.

4. *There exists a nonempty compact set* $\mathcal{B} \subset \mathcal{X}_{0+}$ *such that*
   (i) $\mathcal{B}$ *is invariant under the semiflow* $\{\Phi(t,\cdot)\}_t$.
   (ii) *The subset* $\mathcal{B}$ *attracts the bounded sets of* $\mathcal{X}_{0+}$ *under the semiflow* $\{\Phi(t,\cdot)\}_t$.

We recall that a nonempty set $J \subset \mathcal{X}$ is said to *attract* a nonempty set $B \subset \mathcal{X}$ if $\delta_{\mathcal{X}}(\Phi(t,B), J)) \to 0$ as $t \to +\infty$, where $\delta_{\mathcal{X}}(B, J) = \sup_{u \in B} \inf_{v \in J}\|u - v\|_{\mathcal{X}}$ is a semi-distance on $\mathcal{X}$. Moreover, notice that the notion of *asymptotically smooth* is used here within the terminology of Hale, Lasalle and Slemrod [20]. However, this is strictly equivalent to the notion of *asymptotically compact* within the terminology of Ladyzhenskaya [28]. Finally, we refer to Appendix A for the proof of Theorem 2.1.

2.2. Basic reproduction number and asymptotic behaviour

An equilibrium $(S, i_1(a), i_2(a), i_3(a))$ of system (1.1) is such that $i_j(a) = D_j(a)i_j(0)$, with $j = 1, 2, 3$; and

$$\begin{cases} i_1(0) = \frac{S}{P}\left(i_1(0)\Omega_1 + \varepsilon i_2(0)\Omega_2 + \delta i_3(0)\Omega_3\right), \\ i_2(0) = \Gamma_1 i_1(0), \\ i_3(0) = \Gamma_2 i_2(0), \\ i_1(0) = \Lambda - \mu S, \\ P = S + \sum_{j=1}^3 \overline{D}_j i_j(0). \end{cases}$$

The disease-free equilibrium corresponds to $i_1(0) = i_2(0) = i_3(0) = 0$ and is given by $E^0 = \left(\frac{\Lambda}{\mu}, 0, 0, 0\right)$.

In order to find any endemic equilibria, we first determine the basic reproduction number R_0 using the next generation operator approach [11, 26]. Roughly speaking, we show that the number of new HIV cases $N(t)$ at time t is given by the renewal equation (see Appendix C for details)

$$N(t) = \int_0^\infty K(a)N(t-a)\mathrm{d}a,$$

where

$$K(a) = \beta(a)D_1(a) + \varepsilon \int_0^a D_1(a-\sigma)\gamma_1(a-\sigma)\beta(\sigma)D_2(\sigma)\mathrm{d}\sigma\mathrm{d}a + \delta \int_0^a U(a-\sigma)\beta(\sigma)D_3(\sigma)\mathrm{d}\sigma,$$

and $U(a) = \int_0^a \gamma_1(a-\sigma)D_1(a-\sigma)\gamma_2(\sigma)D_2(\sigma)\mathrm{d}\sigma$. Note that $K(a)$ is the infectiousness of an individual which is infected since time a.

From where we calculate (Appendix C) $R_0 = \int_0^\infty K(a)\mathrm{d}a$; that is

$$R_0 = \Omega_1 + \varepsilon\Omega_2\Gamma_1 + \delta\Omega_3\Gamma_1\Gamma_2,$$

where Ω_j and Γ_j,s are given by (2.1).

Now by taking $i_j(0)$'s positive, by straightforward algebra, the disease-endemic equilibrium $E^* = (S^*, i_1^*(\cdot), i_2^*(\cdot), i_3^*(\cdot))$ is such that

$$S^* = \frac{I_0^*\overline{D}}{R_0 - 1}; \quad i_1^*(\cdot) = I_0^*D_1(\cdot); \quad i_2^*(\cdot) = I_0^*\Gamma_1 D_2(\cdot); \quad i_3^*(\cdot) = I_0^*\Gamma_1\Gamma_2 D_3(\cdot),$$

with $\overline{D} = \overline{D}_1 + \Gamma_1\overline{D}_2 + \Gamma_1\Gamma_2\overline{D}_3$, $I_0^* = \frac{\Lambda(R_0-1)}{\mu\overline{D}+R_0-1}$, and where $\overline{D}_j$ are given by (2.1).

From where, the following proposition summarizes the equilibria of the model.

Proposition 2.2. *Let Assumption 1.1 be satisfied. Then the semiflow $\{\Phi(t,\cdot)\}_t$ provided by Theorem 2.1 has exactly:*

  (i) *One equilibrium, the disease-free equilibrium E^0, if $R_0 \leq 1$.*
  (i) *Two equilibria, the disease-free equilibrium E^0 and disease-endemic equilibrium E^*, if $R_0 > 1$.*

We end this section by providing the following result concerning the asymptotic behaviour of model (1.1) with respect to the R_0.

Theorem 2.3. *Let Assumption 1.1 be satisfied.*

  1. *The disease-free equilibrium E^0 is globally asymptotically stable if $R_0 < 1$ and unstable if $R_0 > 1$.*
  2. *When $R_0 > 1$, then*
     (i) *The disease-endemic equilibrium E^* is locally asymptotically stable for $\{\Phi(t,\cdot)\}_t$.*
     (ii) *If the initial conditions y_0 satisfy $i_{10} = i_{20} = i_{30} = 0$, then the semiflow $\{\Phi(t,y_0)\}_t$ tends to the disease-free equilibrium E^0 for the topology of $\mathcal{X}$.*
     (iii) *If the initial conditions y_0 satisfy $i_{10} + i_{20} + i_{30} > 0$, then the semiflow $\{\Phi(t,y_0)\}_t$ is uniformly persistent in the sense that there exists $\nu > 0$ such that*

$$\liminf_{t\to\infty} \int_0^\infty \beta(a)\left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a)\right)\mathrm{d}a \geq \nu.$$

We refer to Appendix B for the proof of Theorem 2.3.

TABLE 1. Description of the state variables and parameters of the model.

State variables	
$S(t)$	Susceptible individuals at time t
$i_j(t,a)$	HIV individuals stage j at time t, in the stage for duration a (with no ART)
$i_{j,TF}(t,a)$	HIV individuals stage j with ART drop out at time t, in the stage for duration a
$i_{j,TS}(t,a)$	HIV individuals stage j with ART success at time t, in the stage for duration a

Fixed model parameters		
Parameters	Description; Value	Ref
Λ	Entering flux into S; 30	Ass
μ	Exit rate of S; 1/30 PMs	Ass
$\bar{\gamma}_j$	Progression rate from stage j to $j+1$; 1 PMs	Ass
d_j	Death rate of HIV stage $j=1,2$; μ PMs	Ass
d_3	Death rate of HIV late stage; 0.14 per 1000 PYs+ μ	[49]
ρ_0	The rate of infectiousness; 2.48	Ass
ε	Reduced transmission rate of HIV stage 2; 1/26	[21]
δ	Reduced transmission rate of HIV stage 3; 0	Ass

Variable model parameters		
Parameters	Description; Reference value; {Range}	Ref
T_0^1	Duration of HIV stage 1; 2.90; {1.23-6.00 } months	[21]
T_0^2	Duration of HIV stage 2; 120; {108-180 } months	[21]
β_1	Transmission hazard stage 1; 276; {131-509} per 100 PYs	[21]
p_1	Proba. of ART drop out at HIV stage 1;{0-1}	Ass
p_2	Proba. of ART drop out at stage 2 (with no ART at stage 1); {0-1}	Ass
$p_{2,TF}$	Proba. of ART drop out at stage 2 (with ART failure at stage 1); {0-1}	Ass

Cost coefficients in objective functional		
B	Balancing coefficient; {50-80} €	[5]
C_j	Balancing coefficients; {50-80} €	[5]

PMs=person-months; PYs=person-years; Ass=Assumed.

3. MODEL PARAMETERS, TYPICAL MODEL SIMULATION AND SENSITIVITY ANALYSIS

3.1. Setting model parameters

In this section we briefly describe the shape and the values of parameters considered for the simulations of model (1.1). For all simulations, parameters Λ and μ are assumed to be fixed with constant values given in Table 1. We also assume that the disease-induced mortality for HIV stage j, $d_j(a)$, is constant with respect to the duration within the stage (*i.e.* $d_j(a) \equiv d_j$) and the fixed reference value is given in Table 1. Therefore, we more specifically describe the duration-dependent parameters $\beta(a)$ and $\gamma_j(a)$.

3.1.1. Transmission rate

The transmission rate is defined as in [13, 21]. For simplicity, we assume that individuals with AIDS do not substantially contribute to the further spread of HIV (*i.e.*, not risky sexual behavior) such that we can set $\delta \approx 0$. The transmission rate is assumed to be constant for each stage of infection (but only for parameterization purposes, which does not affect our general model formulation). We set $\beta = \rho_0 \times \beta_1$, where $\rho_0 > 0$ is the rate at which an infectious individual infects the susceptible individuals, and β_1 is the transmission hazard. As pointed in [13, 21], HIV is 26 times more infectious during stage 1 than during stage 2, *i.e.*, $\varepsilon = 1/26$. For simulations,

parameters reference values, and ranges are given in Table 1. Further, we set the transmission rate $\rho_0 \simeq 2.48$, such that the maximum value of R_0 in the set of parameters variation is 7 (see [37] for more details on the range of HIV R_0).

3.1.2. Rate of progression from HIV stage j to stage $j + 1$, $\gamma_j(a)$

The parameter γ_j is set to

$$\gamma_j(a) = 0 \text{ if } a < T_0^j \quad \text{and} \quad \bar{\gamma}_j \text{ if } a \geq T_0^j;$$

wherein T_0^j is the duration of the j-stage of HIV infection and $\bar{\gamma}_j = 1$ is the constant progression rate. Infected individuals in stage j remain in that stage for a period of time T_0^j and then progress to stage $j + 1$ at a constant rate $\bar{\gamma}_j$ after the duration of the stage T_0^j. For simulations, the reference values and the range of parameters $\bar{\gamma}_j$ and T_0^j are given in Table 1.

3.2. Typical epidemic dynamics simulated with the model

For all simulations considered in this note, we assume that the initial susceptible population is $S(0) = \Lambda/\mu$ and the initial distribution of infectives are $i_{10}(a) = c_{01}l_{10}(a)$, $i_{20}(a) = c_{02}l_{20}(a)$ and $i_{30}(a) = 0$ for all $a \geq 0$ (in months) with

$$l_{10}(a) = \begin{cases} e^{-\mu a}, a \leq 2.9, \\ e^{-\mu(2a-2.9)}, a > 2.9 \end{cases} \qquad l_{20}(a) = \begin{cases} e^{-\mu a}, a \leq 120, \\ e^{-\mu(2a-120)}, a > 120. \end{cases}$$

By setting $N_0 = S(0) + \|i_{10}\|_{L^1} + \|i_{20}\|_{L^1}$, the constants c_{01} and c_{02} are scaling coefficients given by $c_{01} = \frac{0.025N_0}{100\|l_{10}\|_{L^1}}$ and $c_{02} = \frac{0.025N_0}{100\|l_{20}\|_{L^1}}$ such that the HIV epidemic is initialized with a disease prevalence of $\frac{\|i_{10}\|_{L^1} + \|i_{20}\|_{L^1}}{N_0} \times 100 = 0.05\%$.

Numerical simulations of Figure 1 are based on the reference values of the model parameters defined previously and summarized in Table 1. For those parameter values, $R_0 = 2.55$ and the dynamics of susceptible, $S(t)$, and the total number of HIV stage j, $I_j(t) = \int_0^\infty i_j(t,a)\mathrm{d}a$ with $j = 1, 2, 3$, at time t are given in Figure 1. Furthermore, the total number of HIV late stage I_3 is always be zero in the first 120 months (Fig. 1, right-bottom). This behaviour is explained by the fact that, in Figure 1, the duration of HIV stage 2 is set 120 months and the initial population in HIV late stage is set to zero ($i_{30}(a) \equiv 0$).

3.3. Global sensitivity analysis

Global sensitivity analyses [42] quantify the relative importance of model parameters by partitioning the variance of output variables into those resulting from the main effects of the parameters and their higher-order interactions. Here we study the sensitivity of the HIV late stage $I_3^{tot} = \int_0^T \int_0^\infty i_3(t,a)\mathrm{d}a\mathrm{d}t$ (over the time horizon $T = 300$ months) to the four parameters $\gamma_1(\cdot)$, $\gamma_2(\cdot)$, β_1 and β_2. The range of variation accounting for the known biological variability of the above parameters is assigned in Table 1. Actually, the variability of $\gamma_j(\cdot)$ is determined by the one of T_0^j (see Sect. 3.1). Sensitivity indices are estimated by fitting an ANOVA (Analysis of variance) linear model, including third-order interactions, to the data generated by simulation. Note that this ANOVA linear model fitted very well with 99% of variance explained. The model is implement with MatLab software and the ANOVA analysis with R software (http://www.r-project.org/). Sensitivity analyses indicate that the HIV progression rate from stage 2 to 3 (γ_2) is the most influential factor of the HIV late stage I_3^{tot} (79% of the variance, Fig. 2). The next factor, the transmission rate of stage 1 (β_1) accounting for 8% of the variance, is followed by the HIV progression rate from stage 1 to 2 (γ_1) and the transmission rate of stage 2 (β_2) (<5% of variance explained for each).
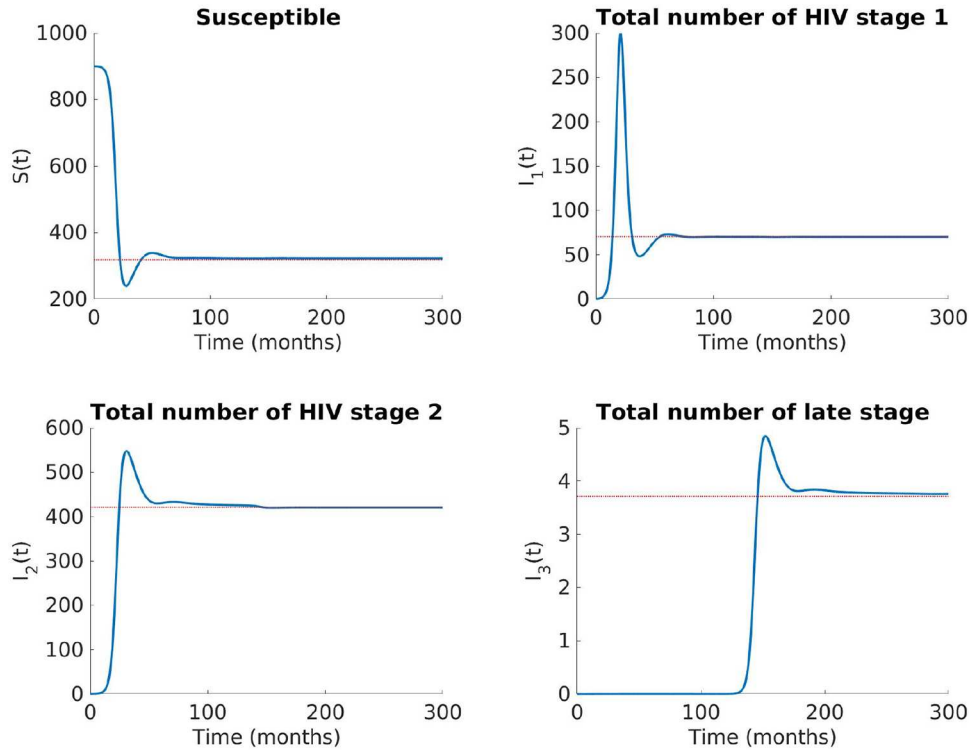
FIGURE 1. Typical epidemics simulated by the model. Parameters of the model are set to their reference values given in Table 1 leading to $R_0 = 2.55$. The figure illustrates the dynamics of susceptible, $S(t)$, and the total number of HIV stage j, $I_j(t) = \int_0^\infty i_j(t,a)\mathrm{d}a$, at time t. For each figure, the dot line represents the endemic equilibrium of the model.
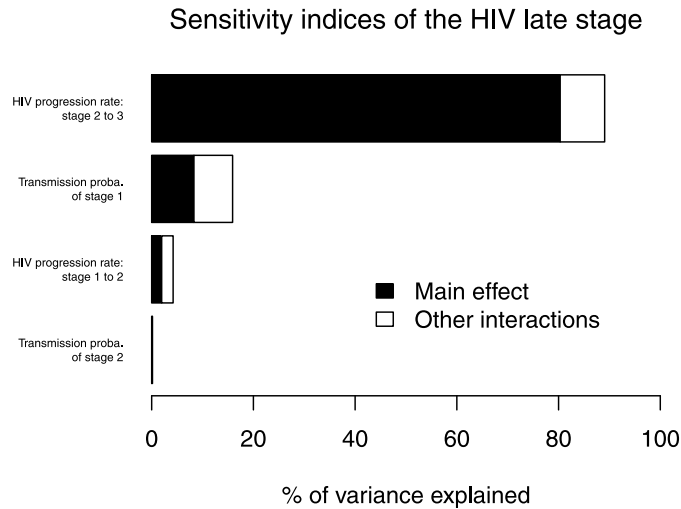

FIGURE 2. Sensitivity indices of the HIV late stage $I_3^{tot} = \int_0^T \int_0^\infty i_3(t,a)\mathrm{d}a\mathrm{d}t$, over $T = 300$ months. The black parts of bars correspond to the main indices (effect of the factor alone) and full bars correspond to total indices (white parts correspond to the effect of the factor in interaction with all other factors).

4. Optimal intervention strategies

Today, more tools are available to prevent HIV such as using condoms the right way every time you have sex or by taking newer HIV prevention medicines such as pre-exposure prophylaxis and post-exposure prophylaxis [6]. But, when living with HIV, up to date, the most important intervention is taking medicines to treat HIV (called ART). Although a cure for HIV does not yet exist, ART can keep healthy for many years if taken consistently and correctly and greatly reduces the chance of transmitting to a partner. Those options are supported by the sensitivity analysis: (i) prevention methods, by the HIV transmission rate and (ii) treatment of HIV by the HIV progression rates (Fig. 2). However, only ART is considered as an intervention strategy in this note.

4.1. Extended model with intervention methods

In addition to the previous state variables, $S(t)$-susceptible, $i_1(t,a)$-HIV stage 1 (who are not under ART) and $i_2(t,a)$-HIV stage 2 (who are not under ART), ART interventions induce four additional state variables: $i_{1,TF}(t,a)$-HIV stage 1 and $i_{2,TF}(t,a)$-HIV stage 2 with ART failure or drop out, as well as $i_{1,TS}(t,a)$-HIV stage 1 and $i_{2,TS}(t,a)$-HIV stage 2 with ART success. We also introduce the probability of treatment drop out: p_1; at HIV stage 1, p_2, at HIV stage 2 (with no ART at HIV stage 1) and $p_{2,TF}$, at HIV stage 2 (with ART drop out at HIV stage 1). This differential infectivity is supported by the fact that individuals who have dropped out of treatment progress through subsequent HIV stages twice as fast as treatment-naive individuals [13]. Therefore, people in class $i_{1,TF}$ (resp. $i_{2,TF}$) progress at rate $\gamma_{1,TF}(a)$ (resp. $\gamma_{2,TF}(a)$) to the HIV stage 2 (resp. late stage). Then, the force of infection and overall progression into subsequent stages write $E_1(t) = \int_0^\infty \beta(a)\left[i_1 + i_{1,TF} + \varepsilon(i_2 + i_{2,TF}) + \delta i_3\right](t,a)\mathrm{d}a$, $E_2(t) = \int_0^\infty \gamma_1(a)i_1(t,a)\mathrm{d}a$, $E_{2,TF}(t) = \int_0^\infty \gamma_{1,TF}(a)i_{1,TF}(t,a)\mathrm{d}a$, and $E_3(t) = \int_0^\infty \left(\gamma_2(a)i_2(t,a) + \gamma_{2,TF}(a)i_{2,TF}(t,a)\right)\mathrm{d}a$. The total population is set to $P(t) = S(t) + \sum_{j=1}^3 \int_0^\infty i_j(t,a)\mathrm{d}a + \sum_{j=1}^2 \int_0^\infty i_{j,TF}(t,a)\mathrm{d}a + \sum_{j=1}^2 \int_0^\infty i_{j,TS}(t,a)\mathrm{d}a$. The transfer diagram of the model becomes:
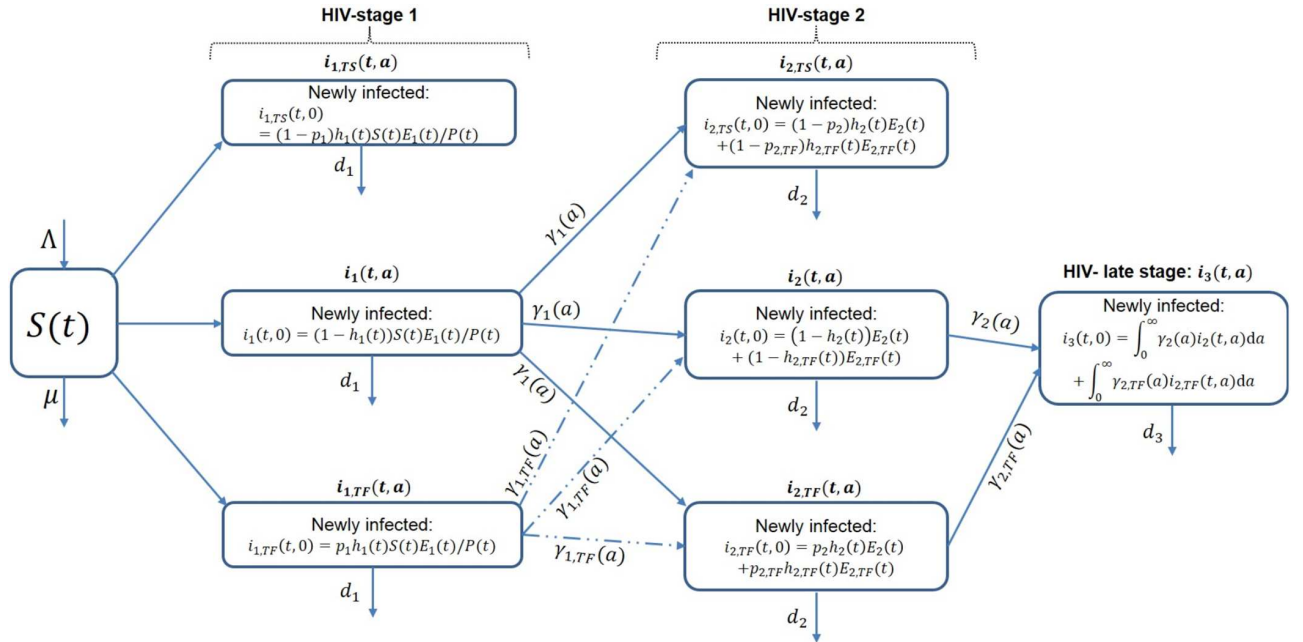
Three intervention strategies, called controls, are included into our initial model (1.1). Controls are represented as functions of time and assigned reasonable upper and lower bounds. At a time t, we define the proportion of individuals: (i) $h_1(t)$; under ART at HIV stage 1, (ii) $h_2(t)$, under ART at HIV stage 2 (which

were not under ART at HIV stage 1) and (iii) $h_{2,TF}(t)$, under ART at HIV stage 2 (with ART drop out at HIV stage 1). Then, the system describing our model with controls writes

$$
\begin{cases}
\dot{S}(t) = \Lambda - \mu S(t) - S(t)E_1(t)/P(t), \\
(\partial_t + \partial_a)\, i_1(t,a) = -\left(\gamma_1(a) + d_1(a)\right) i_1(t,a), \\
(\partial_t + \partial_a)\, i_{1,TF}(t,a) = -\left(\gamma_{1,TF}(a) + d_1(a)\right) i_{1,TF}(t,a), \\
(\partial_t + \partial_a)\, i_{1,TS}(t,a) = -d_1(a) i_{1,TS}(t,a), \\
(\partial_t + \partial_a)\, i_2(t,a) = -\left(\gamma_2(a) + d_2(a)\right) i_2(t,a), \\
(\partial_t + \partial_a)\, i_{2,TF}(t,a) = -\left(\gamma_{2,TF}(a) + d_2(a)\right) i_{2,TF}(t,a), \\
(\partial_t + \partial_a)\, i_{2,TS}(t,a) = -d_2(a) i_{2,TS}(t,a), \\
(\partial_t + \partial_a)\, i_3(t,a) = -d_3(a) i_3(t,a),
\end{cases}
\tag{4.1}
$$

with the boundary conditions

$$
\begin{cases}
i_1(t,0) = \left(1 - h_1(t)\right) S(t)E_1(t)/P(t), \\
i_{1,TF}(t,0) = p_1 h_1(t) S(t)E_1(t)/P(t), \\
i_{1,TS}(t,0) = (1 - p_1) h_1(t) S(t)E_1(t)/P(t), \\
i_2(t,0) = \left(1 - h_2(t)\right) E_2(t) + \left(1 - h_{2,TF}(t)\right) E_{2,TF}(t), \\
i_{2,TF}(t,0) = p_2 h_2(t) E_2(t) + p_{2,TF} h_{2,TF}(t) E_{2,TF}(t), \\
i_{2,TS}(t,0) = (1 - p_2) h_2(t) E_2(t) + (1 - p_{2,TF}) h_{2,TF}(t) E_{2,TF}(t), \\
i_3(t,0) = E_3(t).
\end{cases}
\tag{4.2}
$$

We set $S^h(t)$, and $y^h(t,a) =$
$\left(i_1^h(t,a), i_{1,TF}^h(t,a), i_2^h(t,a), i_{2,TF}^h(t,a), i_3^h(t,a), i_{1,TS}^h(t,a), i_{2,TS}^h(t,a)\right)^T$ the solution of (4.1)–(4.2) associated to the control scheme $h = (h_1, h_2, h_{2,TF})^T$. To illustrate the dependency with respect to h, we also set $P^h(t) = P(t)$, $E_1^h(t) = E_1(t)$, $E_2^h(t) = E_2(t)$, $E_{2,TF}^h(t) = E_{2,TF}(t)$, and $E_3^h(t) = E_3(t)$. Then, problem (4.1)–(4.2) rewrites

$$
\begin{cases}
\dot{S}^h(t) = g_S\left(S^h(t), E_1^h(t)/P^h(t)\right), \\
(\partial_t + \partial_a)\, y^h(t,a) = f(a)y^h(t,a), \\
y^h(t,0) = \phi\left(h(t), E_1^h(t)/P^h(t), E_2^h(t), E_{2,TF}^h(t), E_3^h(t), E_{3,TF}^h(t)\right),
\end{cases}
\tag{4.3}
$$

wherein g_S is given by the right-hand side of (4.1) for the S-compartment; $f(a)y^h$ is the linear operator given by the the right-hand side of (4.1) for the $(i_1, i_{1,TF}, i_2, i_{2,TF}, i_3, i_{1,TS}, i_{2,TS})$-compartment and ϕ is given by the the right-hand side of (4.2) for the $(i_1, i_{1,TF}, i_2, i_{2,TF}, i_3, i_{1,TS}, i_{2,TS})$-compartment. We denote by $\mathcal{V} = \{1, 1.TF, 2, 2.TF, 3, 1.TS, 2.TS\}$ the set of indexes, and system (4.3) is coupled with the initial positive condition

$$
S^h(0) = S_0, i_j^h(0,a) = i_{j0}(a), j \in \mathcal{V}.
\tag{4.4}
$$

The controlled system (4.3) is considered under the below hypotheses:

Assumption 4.1.   1. The control scheme h lies on set

$$
\mathcal{U} = \left\{ h = (h_1, h_2, h_{2,TF})^T \in \left(L^\infty(0, T_f)\right)^3 : 0 \leq h_1(\cdot) \leq h_1^{\max}; 0 \leq h_2(\cdot) \leq h_2^{\max}; 0 \leq h_{2,TF}(\cdot) \leq h_{2,TF}^{\max} \right\},
$$

where $h_j^{\max} \leq 1$ are positive measurable functions, and T^f is the period during which the control is applied.

2. $\Lambda, \mu > 0$; $\beta, \gamma_j(j \in \mathcal{V}), d_1, d_2, d_3 \in L_+^\infty(0, \infty)$ and $d_j(\cdot) \geq \mu$.
3. β and $\gamma_j(j \in \mathcal{V})$ are Lipschitz continuous almost everywhere on $\mathbb{R}_+$.
4. For any $a > 0$, there exists $a_\beta; a_{\gamma_j} > a$, with $j \in \mathcal{V}$; such that β is positive in a neighbourhood of a_β and γ_j is positive in a neighbourhood of a_{γ_j}.

Although our objective is not to make a detailed analysis of the controlled system (4.3)–(4.4), we nevertheless state the following well-posedness result.

Theorem 4.2. *Let Assumption* (4.1) *be satisfied. Then the controlled system* (4.3)–(4.4) *generates a unique globally defined, positive and bounded non-autonomous semiflow.*

The proof of this result is based on a rather standard methodology (*e.g.*, [34, 38]). Here, for instance, we more precisely refer to Theorem 6.2 in [34].

4.2. Optimal control problem

We assume that a successful scheme is one which reduces the progression to the AIDS stage. Therefore, the control scheme is optimal if it minimizes the objective functional

$$
\begin{aligned}
J(h) = \quad & \int_0^{T_f} \int_0^\infty B(a) \left[\gamma_2(a) i_2(t, a) + \gamma_{2,TF}(a) i_{2,TF}(t, a) \right] \mathrm{d}a \mathrm{d}t \\
& + \int_0^{T_f} [C_1 h_1^2(t) + C_2 h_2^2(t) + C_3 h_{2,TF}^2(t)] \mathrm{d}t
\end{aligned}
$$

where B and C_j are balancing coefficients transforming the integral into cost expended over a finite period of T_f months, see Table 1. The first integral, multiplied by B, is the economic losses from individuals progressing into AIDS stage and the second integral represents the costs for the implementation of three controls. Quadratic expressions of controls are included to indicate nonlinear costs potentially arising at high treatment levels.

Our aim is to find h^* satisfying

$$
J(h^*) = \min_{h \in \mathcal{U}} J(h). \tag{4.5}
$$

4.3. The necessary optimality condition

To deal with the necessary optimality conditions, we use some results in Feichtinger *et al.* [16]. For a given solution $S(t), y(t, a) = (i_1, i_{1,TF}, i_2, i_{2,TF}, i_3, i_{1,TS}, i_{2,TS})(t, a)$ and $(E_1(t)/P(t), E_2(t), E_{2,TF}(t), E_3(t))$ of (4.1)–(4.2), we introduce the following adjoint functions $\lambda_S(t), \xi(t, a) = (\lambda_{i_1}, \lambda_{i_{1,TF}}, \lambda_{i_2}, \lambda_{i_{2,TF}}, \lambda_{i_3}, \lambda_{i_{1,TS}}, \lambda_{i_{2,TS}})(t, a)$ and $(\zeta_1(t), \zeta_2(t), \zeta_{2,TF}(t), \zeta_3(t))$.

The following system holds from [16]

$$
\begin{cases}
\dot{\lambda}_S(t) &= \left[E_1(t)/P(t) + \mu\right]\lambda_S(t), \\
(\partial_t + \partial_a)\lambda_{i_1}(t,a) &= \left[\gamma_1(a) + d_1(a)\right]\lambda_{i_1}(t,a) - \beta(a)\zeta_1(t)/P(t) - \gamma_1(a)\zeta_2(t), \\
(\partial_t + \partial_a)\lambda_{i_{1,TF}}(t,a) &= \left[\gamma_{1,TF}(a) + d_1(a)\right]\lambda_{i_{1,TF}}(t,a) - \beta(a)\zeta_1(t)/P(t) - \gamma_{1,TF}(a)\zeta_{2,TF}(t), \\
(\partial_t + \partial_a)\lambda_{i_2}(t,a) &= -B\gamma_2(a) + \left[\gamma_2(a) + d_2(a)\right]\lambda_{i_2}(t,a) - \varepsilon\beta(a)\zeta_1(t)/P(t) - \gamma_2(a)\zeta_3(t), \\
(\partial_t + \partial_a)\lambda_{i_{2,TF}}(t,a) &= -B\gamma_{2,TF}(a) + \left[\gamma_{2,TF}(a) + d_2(a)\right]\lambda_{i_{2,TF}}(t,a) \\
&\quad -\varepsilon\beta(a)\zeta_1(t)/P(t) - \gamma_{2,TF}(a)\zeta_3(t), \\
(\partial_t + \partial_a)\lambda_{i_3}(t,a) &= d_3(a)\lambda_{i_3}(t,a) - \delta\beta(a)\zeta_1(t)/P(t), \\
(\partial_t + \partial_a)\lambda_{i_{1,TS}}(t,a) &= d_1(a)\lambda_{i_{1,TS}}, \\
(\partial_t + \partial_a)\lambda_{i_{2,TS}}(t,a) &= d_2(a)\lambda_{i_{2,TS}}, \\
\zeta_1(t) &= (1 - h_1(t))S(t)\lambda_{i_1}(t,0) + p_1 h_1(t)S(t)\lambda_{i_{1,TF}}(t,0) \\
&\quad +(1 - p_1)h_1(t)S(t)\lambda_{i_{1,TS}}(t,0), \\
\zeta_2(t) &= (1 - h_2(t))\lambda_{i_2}(t,0) + p_2 h_2(t)\lambda_{i_{2,TF}}(t,0) + (1 - p_2)h_2(t)\lambda_{i_{2,TS}}(t,0), \\
\zeta_{2,TF}(t) &= (1 - h_{2,TF}(t))\lambda_{i_2}(t,0) + p_{2,TF}h_{2,TF}(t)\lambda_{i_{2,TF}}(t,0) \\
&\quad +(1 - p_{2,TF})h_{2,TF}(t)\lambda_{i_{2,TS}}(t,0), \\
\zeta_3(t) &= \lambda_{i_3}(t,0),
\end{cases}
\tag{4.6}
$$

with the boundary conditions

$$
\lambda_S(T_f) = 0, \quad \lambda_v(T_f, a) = 0, \text{ for } v \in \{i_1, i_{1,TF}, i_2, i_{2,TF}, i_3, i_{1,TS}, i_{2,TS}\} \text{ and for all } a > 0.
\tag{4.7}
$$

The Hamiltonian of System (4.1)–(4.2) is given by

$$
H(t,h) = \xi(t,0) \cdot \phi\left(h, E_1/P, E_2, E_{2,TF}, E_3\right) + \lambda_S(t)g_S(t) + L_0(t) + \int_0^\infty L(t,a)\mathrm{d}a,
$$

where L and L_0 are the first and second integrand of J respectively.

Moreover, if h^* is a solution of (4.5), then it is characterized by

$$
\begin{aligned}
h_1^*(t) &= \max(0, \min(\hat{h}_1(t), h_1^{\max})), \\
h_2^*(t) &= \max(0, \min(\hat{h}_2(t), h_2^{\max})), \\
h_{2,TF}^*(t) &= \max(0, \min(\hat{h}_{2,TF}(t), h_{2,TF}^{\max})),
\end{aligned}
\tag{4.8}
$$

wherein

$$
\begin{aligned}
\hat{h}_1(t) &= \frac{\left(\lambda_{i_1}(t,0) - p_1\lambda_{i_{1,TF}}(t,0) - (1 - p_1)\lambda_{i_{1,TS}}(t,0)\right)S(t)E_1(t)/P(t)}{2C_1}, \\
\hat{h}_2(t) &= \frac{\left(\lambda_{i_2}(t,0) - p_2\lambda_{i_{2,TF}}(t,0) - (1 - p_2)\lambda_{i_{2,TS}}(t,0)\right)E_2(t)}{2C_2}, \\
\hat{h}_{2,TF}(t) &= \frac{\left(\lambda_{i_2}(t,0) - p_{2,TF}\lambda_{i_{2,TF}}(t,0) - (1 - p_{2,TF})\lambda_{i_{2,TS}}(t,0)\right)E_{2,TF}(t)}{2C_3}.
\end{aligned}
$$

The value function $\hat{h}_1$ is obtained from equation $\dfrac{\partial H}{\partial h_1} = 0$, using the boundary condition whenever $0 \leq h_1^*(t) \leq h_1^{\max}$, and similarly for $\hat{h}_2$ and $\hat{h}_{2,TF}$.

The existence and uniqueness of a solution to (4.5) are given by the following result

Theorem 4.3. *If $\frac{T_f}{2}\sum_{j=1}^{3}\frac{1}{C_j}$ is sufficiently small, there exists one and only one optimal controller h^* in $\mathcal{U}$ minimizing $J(h)$.*

The proof of the above theorem is given in Appendix D. Let us emphasize that, since the functional J is not weakly lower semi-continuous on L^1, therefore, we might not expect existence for the problem (4.5) in general. This is a common feature of nonlinear optimal control problems (*e.g.,* see [4]) and one can prove the existence for some time-interval using the Ekeland variational principle [14].

The state system (4.1)–(4.2) and the adjoint system (4.6)–(4.7) together with the control characterization system (4.8) form the optimality system to be solved numerically. Those systems constitute a feedback system in the sense that: (i) a given controller is used as the input of the state system, (ii) the controller and the output of the state system are used as the input of the adjoint system, (iii) the outputs of the state and adjoint systems are used to update the controller through the control characterization system, and (iv) the updated controller becomes the input of the state system, creating a "closed loop" system. Finally, since the state equations have initial conditions and the adjoint equations have final time conditions, we cannot solve the optimality system directly by only sweeping forward in time. Thus, an iterative algorithm, "forward-backward sweep method", is used (see [29]).

5. Discussion

For all simulations considered in this section, in addition to the parameter reference values given by Table 1, the number of months for the strategy deployment is $T_f = 420$ (35 years). Moreover, the progression rate in classes $i_{1,TF}$ and $i_{2,TF}$, $\gamma_{1,TF}(a)$ and $\gamma_{2,TF}(a)$, have the same shape as $\gamma_1(a)$ and $\gamma_2(a)$ defined in Section 3.1. However, the duration of each stage $i_{1,TF}$ and $i_{2,TF}$ is $T_0^1/2$ and $T_0^2/2$ respectively [13]. At the beginning of intervention strategies, initial condition of state variables i_1, i_2 and i_3 is the same as in Section 3.2. Furthermore, we set $i_{1,TS}(0,a) = i_{1,TF}(0,a) = i_{2,TS}(0,a) = i_{2,TF}(0,a) = 0$ for all $a \geq 0$.

5.1. Performance of control strategies

The performance is estimated by assessing the total number of AIDS cases during T months relatively to AIDS cases without any strategy. Formally, the performance of the intervention strategy $h = (h_1, h_2, h_{2,TF})$ is

$$\Delta^h(P) = \frac{\int_0^T \int_0^\infty i_3(t,a)\big|_h \mathrm{d}a\mathrm{d}t}{\int_0^T \int_0^\infty i_3(t,a)\big|_{h=0} \mathrm{d}a\mathrm{d}t},$$

wherein $P = (p_1, p_2, p_{2,TF})$. $\Delta^h(P)$ provides an estimate of the number of AIDS obtained with the control h ($\int\int i_3(t,a)\big|_h \mathrm{d}a\mathrm{d}t$) relatively to the number of AIDS that would have been obtained without any control ($\int\int i_3(t,a)\big|_{h=0}\mathrm{d}a\mathrm{d}t$). For example, $\Delta^h(P) = 0.3$ indicates that the number of AIDS with control is $1/0.3 \approx 3$ times smaller than without control. However, a value of $\Delta^h(P) > 1$ indicates that the intervention strategy has a negative impact on the epidemic outbreak.

Figure 3 illustrates the effect of intervention strategies on the number of AIDS, $I_3(t) = \int_0^\infty i_3(t,a)\mathrm{d}a$. Parameters of the model are set to their reference values given in Table 1 leading to $R_0 = 2.55$ (without controls) and we also set $p_1 = p_2 = p_{2,TF} = 10\%$. Optimal three-part intervention strategies, h_1, h_2 and $h_{2,TF}$, provide considerable reductions in the severity of the projected outbreaks (Fig. 3A and B). Indeed, the performance of controls is $\Delta^h(P) \approx 0.018$, meaning that the total number of AIDS cases with control is 55 times smaller than without control.

5.2. Effect of intervention only at HIV stage 1 or 2

Notice that the optimal control problem can be formulated to find the optimal strategy of each HIV stage intervention method when used alone. Moreover, without any intervention at HIV stage 1, the only plausible
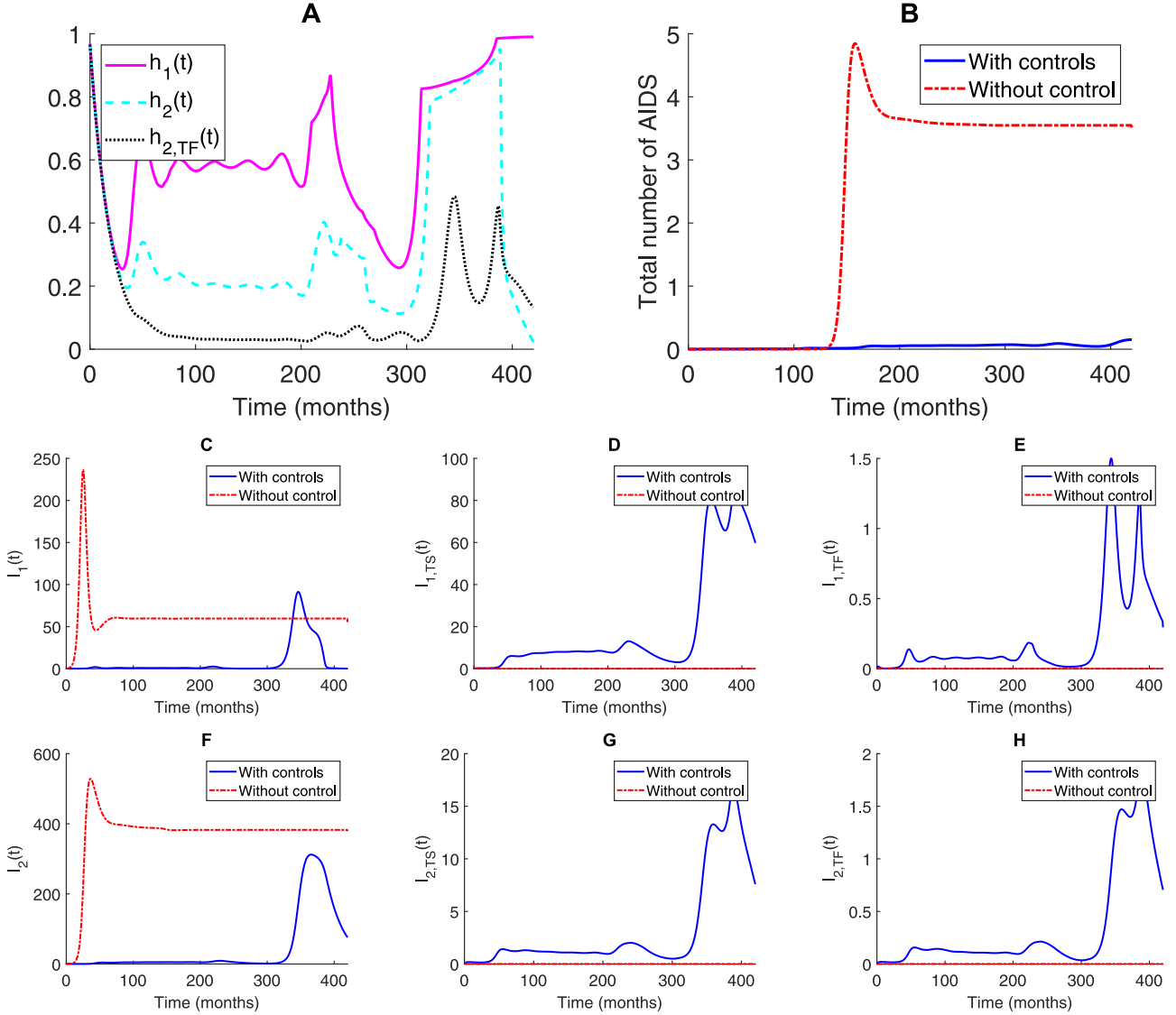
FIGURE 3. Effect of intervention strategies on the number of AIDS. (A) Intervention strategies h_1, h_2 and $h_{2,TF}$ with respect to the time t. (B) The dynamics of the total number of AIDS cases, $I_3(t) = \int_0^\infty i_3(t,a)\mathrm{d}a$, with and without controls. (C-H) The dynamics of the total number of HIV stage j, $I_j(t) = \int_0^\infty i_j(t,a)\mathrm{d}a$, $j \in \{1, 1.TS, 1.TF, 2, 2.TS, 2.TF\}$, with and without controls. Parameters of the model are set to their reference values given in Table 1 leading to $R_0 = 2.55$ (without controls). We also set $p_1 = p_2 = p_{2,TF} = 10\%$.

intervention at stage 2 is h_2. The control h_1 alone has a significant effect on the epidemic outbreaks with a performance $\Delta^{h_1}(p_1) \leq 0.2$ when the probability of treatment drop out at HIV stage 1 p_1 ranges from 0 to 80% in the host population. In other words, for $p_1 \leq 0.8$, the total number of AIDS cases with controls h_1 is at least, 5 times smaller than without control. See Figure 4. The control h_2 alone has a significant effect on the epidemic outbreaks only when the probability of treatment drop out at HIV stage 2, p_2 (with not ART at stage 1) is very small ($p_2 < 0.1$). The performance $\Delta^{h_2}(p_2)$ increases linearly with the value of p_2 and crosses the unity
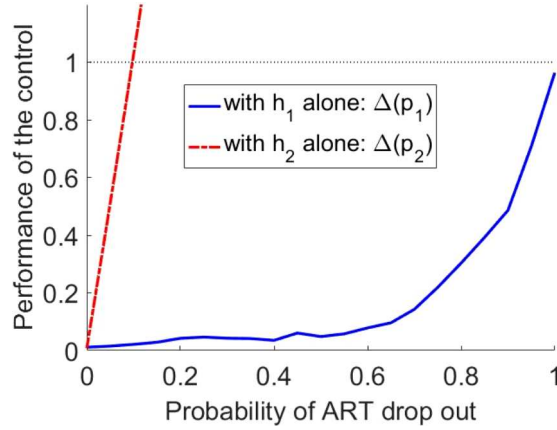
FIGURE 4. Performance of HIV intervention strategy when used alone. (Dotted line) Performance of HIV stage 1 intervention alone $\Delta^{h_1}(p_1)$ with respect to the probability of ART drop out at stage 1, p_1. (Solid line) Performance of HIV stage 2 intervention alone $\Delta^{h_2}(p_2)$ with respect to the probability of ART drop out at stage 2, p_2 (with no ART at stage 1).
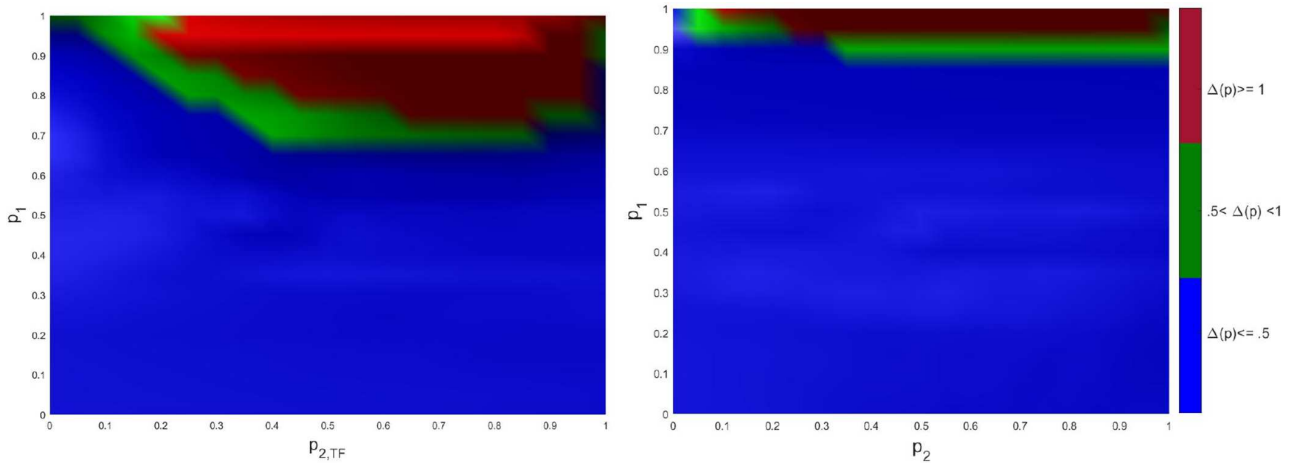

FIGURE 5. Combined effect of intervention at HIV stages 1 and 2. (*Left*) The performance $\Delta^{h_1,h_{2,TF}}(p_1, p_{2,TF})$ of controls h_1 and $h_{2,TF}$ with respect to probabilities of treatment drop out at HIV stage 1, p_1 and treatment drop out at HIV stage 2 (with ART failure at stage 1), $p_{2,TF}$. (*Right*) The performance $\Delta^{h_1,h_2}(p_1, p_2)$ of controls h_1 and h_2 with respect to probabilities of treatment drop out at HIV stage 1, p_1 and treatment drop out at HIV stage 2 (with no ART at stage 1), p_2.

around $p_2 = 10\%$, after what the intervention strategy will have a negative effect on the epidemic outbreaks in the host population, Figure 4. Therefore, before introducing ART, investigations should be addressed to know whether the host population is well sensibilized on ART treatment or not. These investigations will probably help in reducing the probability of treatment drop out.

5.3. Combined effect of intervention at both HIV stages

Controls h_1 and $h_{2,TF}$ have a significant effect on the disease outbreak with the performance of both controls $\Delta^{h_1,h_{2,TF}}(p_1, p_{2,TF}) < .5$, when the probability of treatment drop out at HIV stage 1, $p_1 \leq 60\%$ and what ever

the probability of treatment drop out at HIV stage 2 (with ART failure at stage 1), $p_{2,TF} \in (0,1)$. See Figure 5 (Left). However, for high values of p_1 ($p_1 > .6$) and for $p_{2,TF} > .2$, both controls have a negative effect on the epidemics outbreak with the performance $\Delta^{h_1,h_{2,TF}}(p_1, p_{2,TF}) \geq 1$, see Figure 5 (Left). These configurations are quite similar with the combined effect of controls h_1 and h_2, see Figure 5 (Right). However, notice that controls h_1 and h_2 performed better than controls h_1 and $h_{2,TF}$, even for values of p_1 up to 80%.

5.4. Populations heterogeneity

The lack of various mixing patterns in the host population is a potential limitation in the model formulation. Indeed, sexual behaviour influences HIV transmission dynamics in various ways and studies showed the impact of changes in risk behaviour of individuals over time on HIV transmission dynamics. It is shown that heterogeneity in contact rates among different subgroups of the population and certain forms of mixing (*e.g.*, proportionate mixing) does not substantially impact the HIV transmission dynamics [27]. However, the variance in both the rate of partner change and mixing between subpopulations with different HIV risk levels have a large influence on endemic prevalence before the introduction of ART and on possible long term effectiveness of ART [41].

CONFLICT OF INTEREST

We declare we have no competing interest.

APPENDIX A. PROOF OF THEOREM 2.1

It is easy to check that the operator A is a Hille-Yosida operator. Then standard results apply to provide the existence and uniqueness of a mild solution to (1.1) (we refer to [31, 46, 48] for more details). The Volterra formulation is also standard and we refer to [25, 50] for more details.

For estimate 2., let $\varphi_0 \in \mathcal{X}_{0+}$; then adding up the S equation together with the i_j equations of (1.1) yields

$$\dot{P}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(S(t) + \sum_{j=1}^{3}\int_0^\infty i_j(t,a)\mathrm{d}a\right) \leq \Lambda - \mu P(t).$$

From where one deduces estimate 2.

The bounded dissipativity of the semiflow $\{\Phi(t,\cdot)\}_t$ is a direct consequence of estimate 2. It remains to prove the asymptotic smoothness. For that end, let B be a forward invariant bounded subset of $\mathcal{X}_{0+}$. According to the results in [43] it is sufficient to show that the semiflow is asymptotically compact on B. We first claim that

Claim A.1. *Let Assumption 1.1 be satisfied. Then, functions E_j's are Lipschitz continuous on $\mathbb{R}_+$.*

Therefore, let us consider a sequence of solutions $(S^n, i_1^n, i_2^n, i_3^n)_p$ that is equibounded in $\mathcal{X}_{0+}$ and a sequence $\{t_n\}_p$ such that $t_n \to +\infty$. Let $I_j^n(t) = \int_0^\infty i_j^n(t,\sigma)\mathrm{d}\sigma$, $P^n(t) = S^n(t) + \sum_{j=1}^3 I_j^n(t)$ and E_j^n (defined by (2.4)); with $j = 1,2,3$; the corresponding sequences. Since S^n, P^n, I_j^n's and E_j^n's are uniformly bounded in the Lipschitz norm, the Arzela-Ascoli theorem implies that, possibly along a subsequence, one may assume that $S^n(t+t_n) \to \tilde{S}$, $P^n(t+t_n) \to \tilde{P}$ and $E_j^n(t+t_n) \to \tilde{E}_j$ locally uniformly for for $t \in \mathbb{R}$. It remains to deal with the sequences $\{i_j^n\}_n$ with $j = 1,2,3$. Denoting by $\tilde{i}_1^n(t,\cdot) = i_1^n(t+t_n,\cdot)$ and using the Volterra formulation (2.3) it comes

$$\tilde{i}_1^n(t,a) = \begin{cases} i_{10}(a-t+t_n)\frac{D_1(a)}{D_1(a-t+t_n)}, & \text{for } t+t_n < a, \\ \frac{S^n(t-a+t_n)}{P^n(t-a+t_n)}E_1^n(t-a+t_n)D_1(a), & \text{for } t+t_n \geq a. \end{cases}$$

Since $\frac{S^n(t-a+t_n)}{P^n(t-a+t_n)}E_1^n(t-a+t_n)D_1(a)$ converges locally uniformly towards the function $\frac{\tilde{S}(t-a)}{\tilde{P}(t-a)}\tilde{E}_1(t-a)D_1(a)$ as $t_n \to +\infty$, we easily conclude that

$$i_1^n(t_n, \cdot) = i_1^{\tilde{n}}(0, \cdot) \to \frac{\tilde{S}(t-\cdot)}{\tilde{P}(t-\cdot)}\tilde{E}_1(t-\cdot)D_1(\cdot) \text{ in } L^1(0, \infty, \mathbb{R}).$$

Similarly, we also find that

$$i_j^n(t_n, \cdot) = i_j^{\tilde{n}}(0, \cdot) \to \tilde{E}_j(t-\cdot)D_j(\cdot) \text{ in } L^1(0, \infty, \mathbb{R}), \text{ and for } j = 2, 3.$$

Item 3. follows.

For item 4. of the theorem, items 2. and 3. show that Φ is point dissipative, eventually bounded on bounded sets, and asymptotically smooth. Thus, item 4. follows from Theorem 2.33 of [45].

To complete the proof of the theorem, it remains to proof Claim A.1.

Proof of Claim A.1. Let $t \in \mathbb{R}^+$ and $\eta > 0$. Recalling (2.4) and setting $B_1(t) = \int_0^\infty \beta(a)i_1(t, a)\mathrm{d}a$, $B_2(t) = \varepsilon \int_0^\infty \beta(a)i_2(t, a)\mathrm{d}a$ and $B_3(t) = \delta \int_0^\infty \beta(a)i_3(t, a)\mathrm{d}a$, it comes

$$\begin{aligned} B_1(t+\eta) - B_1(t) &= \int_0^\eta \beta(a)i_1(t+\eta, a)\mathrm{d}a + \int_\eta^\infty \beta(a)i_1(t+\eta, a)\mathrm{d}a - \int_0^\infty \beta(a)i_1(t, a)\mathrm{d}a \\ &= \int_0^\eta \beta(a)i_1(t+\eta-a, 0)D_1(a)\mathrm{d}a + \int_\eta^\infty \beta(a)i_1(t+\eta, a)\mathrm{d}a - \int_0^\infty \beta(a)i_1(t, a)\mathrm{d}a. \end{aligned}$$

Since the semiflow Φ is bounded and by Assumption 1.1 (item 2.), we can find $C > 0$ such that

$$B_1(t+\eta) - B_1(t) \leq C\|\beta\|_\infty^2 \eta + \int_0^\infty \beta(a+\eta)i_1(t+\eta, a+\eta)\mathrm{d}a - \int_0^\infty \beta(a)i_1(t, a)\mathrm{d}a.$$

Then, recalling (2.3) and combining the integrals, we write

$$\begin{aligned} B_1(t+\eta) - B_1(t) \leq &C\|\beta\|_\infty^2 \eta + \int_0^\infty \beta(a+\eta)\left(e^{-\int_a^{a+\eta}(\gamma_1(\sigma)+d_1(\sigma))\mathrm{d}\sigma} - 1\right)i_1(t, a)\mathrm{d}a \\ &+ \int_0^\infty (\beta(a+\eta) - \beta(a))\,i_1(t, a)\mathrm{d}a. \end{aligned}$$

Again by Assumption 1.1 (item 2.), we have $1 \geq e^{-\int_a^{a+\eta}(\gamma_1(\sigma)+d_1(\sigma))\mathrm{d}\sigma} \geq e^{-(\|\gamma_1\|_\infty+\|d_1\|_\infty)\eta} \geq 1 - (\|\gamma_1\|_\infty + \|d_1\|_\infty)\eta$. Therefore, $\beta(a+\eta)\left|e^{-\int_a^{a+\eta}(\gamma_1(\sigma)+d_1(\sigma))\mathrm{d}\sigma} - 1\right| \leq \|\beta\|_\infty(\|\gamma_1\|_\infty + \|d_1\|_\infty)\eta$ and since the semiflow Φ is bounded we can find a positive constant C such that

$$\begin{aligned} B_1(t+\eta) - B_1(t) \leq &C\|\beta\|_\infty^2 \eta + C\|\beta\|_\infty(\|\gamma_1\|_\infty + \|d_1\|_\infty)\eta \\ &+ \int_0^\infty (\beta(a+\eta) - \beta(a))\,i_1(t, a)\mathrm{d}a. \end{aligned}$$

Next, using successively Assumption 1.1 (namely, the Lipschitzianity of β) and the boundedness of the semiflow, we find a positive constant C such that

$$\int_0^\infty |\beta(a+\eta) - \beta(a)|\,i_1(t, a)\mathrm{d}a \leq C\eta \int_0^\infty i_1(t, a)\mathrm{d}a \leq C\eta.$$

From where, we find $C_1 > 0$ such that

$$|B_1(t + \eta) - B_1(t)| \le C_1\eta.$$

Using the same argument, we find $C_2, C_3 > 0$ such that

$$|B_2(t + \eta) - B_2(t)| \le C_2\eta, \text{ and } |B_3(t + \eta) - B_3(t)| \le C_3\eta.$$

Since $E_1 = B_1 + B_2 + B_3$, combining the previous inequalities, it follows that E_1 is Lipschitz on $\mathbb{R}_+$. Similarly, E_2 and E_3 are also Lipschitz and this ends the proof of the claim. $\qquad\square$

APPENDIX B. PROOF OF THEOREM 2.3

First, let us introduce some useful technical materials by establishing some properties of the complete solutions of system (1.1).

B.1 Technical materials

The first result deals with spectral properties of the linearized semiflow Φ at a given equilibrium point $\varphi^* = (v^*, 0_{\mathbb{R}^3}, u_1^*, u_2^*, u_3^*) \in \mathcal{X}_{0+}$. The linearized system at φ^* reads

$$\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = (A + F^*)\,\varphi(t),$$

where F^* is a linear bounded operator given by

$$F^*\left(v, 0_{\mathbb{R}^3}, u_1, u_2, u_3\right)^T = \left(\; -\mathcal{W}^*, \mathcal{W}^*, \int_0^\infty \gamma_1(a)u_1(a)\mathrm{d}a, \int_0^\infty \gamma_2(a)u_2(a)\mathrm{d}a, 0, 0, 0 \;\right)^T,$$

wherein

$$\mathcal{W}^* = \frac{1}{\left(v^* + \overline{A}^*\right)^2}\left[v\overline{A}^*\overline{B}^* - v^*\overline{B}^* \int_0^\infty \sum_{j=1}^3 u_j(a)\mathrm{d}a \right.$$
$$\left. + v^*\left(v^* + \overline{A}^*\right)\int_0^\infty \beta(a)\left(u_1(a) + \varepsilon u_2(a) + \delta u_3(a)\right)\mathrm{d}a\right],$$

and

$$\overline{A}^* = \int_0^\infty \sum_{j=1}^3 u_j^*(a)\mathrm{d}a,$$
$$\overline{B}^* = \int_0^\infty \beta(a)\left(u_1^*(a) + \varepsilon u_2^*(a) + \delta u_3^*(a)\right)\mathrm{d}a.$$

Lemma B.1. *Let us set $\Sigma = \{\lambda \in \mathbb{C} : \Re(\lambda) > -\mu\}$. The spectrum $\sigma(A + F^*) \cap \Sigma$ consists of a point spectrum and one has*

$$\sigma(A + F^*) \cap \Sigma = \{\lambda \in \Sigma : \chi(\lambda, \varphi^*) = 0\}$$

where the function $\chi(\cdot, \varphi^*) : \Sigma \to \mathbb{C}$ *is defined by*

$$\chi(\cdot, \varphi^*) = 1 - \frac{\mathcal{Q}^*(\cdot)(\cdot + \mu)}{\overline{A}^*\overline{B}^* + \left(v^* + \overline{A}^*\right)^2 (\cdot + \mu)}, \tag{B.1}$$

and wherein

$$\mathcal{Q}^*(\lambda) = - v^*\overline{B}^* \left(D_1^\lambda + D_2^\lambda\Gamma_1(\lambda) + D_3^\lambda\Gamma_1(\lambda)\Gamma_2(\lambda)\right)$$
$$+ v^* \left(v^* + \overline{A}^*\right) \left(\Omega_1(\lambda) + \varepsilon\Omega_2(\lambda)\Gamma_1(\lambda) + \delta\Omega_3(\lambda)\Gamma_1(\lambda)\Gamma_2(\lambda)\right),$$
$$\Omega_j(\lambda) = \int_0^\infty \beta(a)D_j(a)e^{-\lambda a}\mathrm{d}a, \quad \Gamma_j(\lambda) = \int_0^\infty \gamma_j(a)D_j(a)e^{-\lambda a}\mathrm{d}a,$$
$$D_j^\lambda = \int_0^\infty D_j(a)e^{-\lambda a}\mathrm{d}a.$$

Proof. Let us denote by $A_0 : D(A_0) \subset \mathcal{X}_0 \to \mathcal{X}_0$ the part of A in $\mathcal{X}_0 = \overline{D(A)}$, which is defined by

$$A_0\varphi = A\varphi, \quad \forall \varphi \in D(A_0) = \{\varphi \in D(A) : A\varphi \in \overline{D(A)}\}.$$

Then it is the infinitesimal generator of a C_0-semigroup on $\mathcal{X}_0$ denoted by $\{T_{A_0}(t)\}_t$. We can easily check that the essential growth rate of this semigroup satisfies $\omega_{0,\text{ess}}(A_0) \leq -\mu$. Since operator F^* is compact, results in [12, 47] apply and ensure that the essential growth rate of $\{T_{(A+F^*)_0}(t)\}_t$, the C_0-semigroup generated by the part of $A + F^*$ in $\mathcal{X}_0$ is such that $\omega_{0,\text{ess}}(A + F^*)_0 \leq -\mu$. Applying the results in [32], the latter inequality ensures that $\sigma(A + F^*) \cap \Sigma$ is only composed of a point spectrum of $(A + F^*)$. The derivation of the characteristic equation is standard and we refer to [7, 33]. Indeed, let us first notice that A is a Hille-Yosida operator on $\mathcal{X}$, because for $z > -\mu$ one has $\|(zI - A)^{-1}\| \leq (z + \mu)^{-1}$. Therefore, for $\lambda \in \mathbb{C}$ such that $\Re(\lambda) > -\mu$ it is easily checked that $[\lambda I - (A + F^*)]^{-1} = (\lambda I - A)^{-1}[I - F^*(\lambda I - A)^{-1}]^{-1}$. That is to say

$$\lambda \in \sigma(A + F^*) \Leftrightarrow 1 \in \sigma(F^*(\lambda I - A)^{-1}).$$

We also have

$$[I - F^*(\lambda I - A)^{-1}](\alpha_1, \alpha_2, \alpha_3, \alpha_4, w_1, w_2, w_3)^T = (v, 0_{\mathbb{R}^3}, u_1, u_2, u_3)^T \Leftrightarrow$$
$$\begin{cases} \alpha_1 + \widetilde{\mathcal{W}}^* = v, \\ \alpha_2 - \widetilde{\mathcal{W}}^* = 0, \\ \alpha_3 - \int_0^\infty \gamma_1\tilde{w}_1(a)\mathrm{d}a = 0, \\ \alpha_4 - \int_0^\infty \gamma_2\tilde{w}_2(a)\mathrm{d}a = 0, \\ w_1 = u_1; w_2 = u_2; w_3 = u_3, \end{cases} \tag{B.2}$$

where $\tilde{w}_j(a) = \alpha_{j+1}D_j(a)e^{-\lambda a} + g_j(\lambda, a)$; $g_j(\lambda, a) = \int_0^a u_j(s)\frac{D_j(a)}{D_j(s)}e^{-\lambda(a-s)}\mathrm{d}s$, and

$$\widetilde{\mathcal{W}}^* = \frac{\alpha_1\overline{A}^*\overline{B}^*}{\left(v^* + \overline{A}^*\right)^2(\lambda + \mu)} + \frac{1}{\left(v^* + \overline{A}^*\right)^2}\left[-v^*\overline{B}^*\int_0^\infty\sum_{j=1}^3 \tilde{w}_j(a)\mathrm{d}a\right.$$
$$\left. + v^*\left(v^* + \overline{A}^*\right)\int_0^\infty \beta(a)\left(\tilde{w}_1(a) + \varepsilon\tilde{w}_2(a) + \delta\tilde{w}_3(a)\right)\mathrm{d}a\right].$$

First, from (B.2), we have

$$\alpha_3 = \alpha_2 \Gamma_1(\lambda) + \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a,$$

$$\alpha_4 = \alpha_2 \Gamma_1(\lambda) \Gamma_2(\lambda) + \Gamma_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \gamma_2(a) g_2(\lambda, a) \mathrm{d}a.$$

From where

$$\int_0^\infty \sum_{j=1}^3 \tilde{w}_j(a) \mathrm{d}a = \left(D_1^\lambda + D_2^\lambda \Gamma_1(\lambda) + D_3^\lambda \Gamma_1(\lambda) \Gamma_2(\lambda) \right) \alpha_2$$

$$+ D_3^\lambda \left(\Gamma_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \gamma_2(a) g_2(\lambda, a) \mathrm{d}a \right)$$

$$+ D_2^\lambda \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \sum_{j=1}^3 g_j(\lambda, a) \mathrm{d}a,$$

and

$$\int_0^\infty \beta(a) \left(\tilde{w}_1(a) + \varepsilon \tilde{w}_2(a) + \delta \tilde{w}_3(a) \right) \mathrm{d}a = \left(\Omega_1(\lambda) + \varepsilon \Omega_2(\lambda) \Gamma_1(\lambda) + \delta \Omega_3(\lambda) \Gamma_1(\lambda) \Gamma_2(\lambda) \right) \alpha_2$$

$$+ \delta \Omega_3(\lambda) \left(\Gamma_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \varepsilon \int_0^\infty \gamma_2(a) g_2(\lambda, a) \mathrm{d}a \right)$$

$$+ \varepsilon \Omega_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \beta(a) \left(g_1(\lambda, a) + \varepsilon g_2(\lambda, a) + \delta g_3(\lambda, a) \right) \mathrm{d}a.$$

From the tow previous equality, $\widetilde{\mathcal{W}}^*$ rewrites

$$\widetilde{\mathcal{W}}^* = \frac{\alpha_1 \overline{A}^* \overline{B}^*}{\left(v^* + \overline{A}^* \right)^2 (\lambda + \mu)} + \frac{1}{\left(v^* + \overline{A}^* \right)^2} \left[\mathcal{Q}^* \alpha_2 + R^*(\lambda, u) \right],$$

with

$$R^*(\lambda, u) = -v^* \overline{B}^* D_3^\lambda \left(\Gamma_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \gamma_2(a) g_2(\lambda, a) \mathrm{d}a \right)$$

$$- v^* \overline{B}^* D_2^\lambda \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \sum_{j=1}^3 g_j(\lambda, a) \mathrm{d}a$$

$$+ v^* \left(v^* + \overline{A}^* \right) \delta \Omega_3(\lambda) \left(\Gamma_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \varepsilon \int_0^\infty \gamma_2(a) g_2(\lambda, a) \mathrm{d}a \right)$$

$$+ v^* \left(v^* + \overline{A}^* \right) \varepsilon \Omega_2(\lambda) \int_0^\infty \gamma_1(a) g_1(\lambda, a) \mathrm{d}a + \int_0^\infty \beta(a) \left(g_1(\lambda, a) + \varepsilon g_2(\lambda, a) + \delta g_3(\lambda, a) \right) \mathrm{d}a.$$

From the two first equations of (B.2), a straightforward computation gives

$$\alpha_1 = -\left(1 + \frac{\overline{A}^*\overline{B}^*}{\left(v^* + \overline{A}^*\right)^2(\lambda + \mu)}\right)^{-1}\frac{\mathcal{Q}^*(\lambda)\alpha_2}{\left(v^* + \overline{A}^*\right)^2} + G(\lambda, v, u),$$

$$\alpha_2\left[1 - \frac{\mathcal{Q}^*(\lambda)(\lambda + \mu)}{\overline{A}^*\overline{B}^* + \left(v^* + \overline{A}^*\right)^2(\lambda + \mu)}\right] = \frac{R^*(\lambda, u)}{\left(v^* + \overline{A}^*\right)^2} + \frac{\overline{A}^*\overline{B}^*G(\lambda, v, u)}{\left(v^* + \overline{A}^*\right)^2(\lambda + \mu)}$$

with

$$G(\lambda, v, u) = \left(1 + \frac{\overline{A}^*\overline{B}^*}{\left(v^* + \overline{A}^*\right)^2(\lambda + \mu)}\right)^{-1}\left(v - \frac{R^*(\lambda, u)}{\left(v^* + \overline{A}^*\right)^2}\right).$$

By setting

$$\chi(\lambda, \varphi^*) = 1 - \frac{\mathcal{Q}^*(\lambda)(\lambda + \mu)}{\overline{A}^*\overline{B}^* + \left(v^* + \overline{A}^*\right)^2(\lambda + \mu)},$$

we can then isolate α_2 (and then $\alpha_1, \alpha_3, \alpha_4$) if and only if $\chi(\lambda, \varphi^*) \neq 0$. $\qquad\square$

The next result relies on some properties of the complete solutions of system (1.1).

Lemma B.2. *Set $y_0 = (S_0, i_{10}, i_{20}, i_{30})$ and let $\{y(t) = (S(t), i_1(t, \cdot), i_2(t, \cdot), i_3(t, \cdot))\}_{t \in \mathbb{R}} \subset \mathcal{X}_{0+}$ a complete solution of (1.1) passing through y_0. Then, $S(t)$ is strictly positive for all t and either E_1 is identically zero or $E_1(t)$ is positive for all t.*

Proof. Assume that there exists $t_1 \in \mathbb{R}$ such that $S(t_1) = 0$. Then, the S-equation of (1.1) gives $\dot{S}(t_1) > 0$ meaning that we can find $\delta > 0$ sufficiently small such that $S(t_1 - \delta) < 0$. A contradiction with the fact that the total trajectory $y(t)$ lies in $\mathcal{X}_{0+}$ in for all $t \in \mathbb{R}$.

Next, let us notice that since y is a complete solution, it comes from the Volterra formulation

$$i_1(t, a) = \frac{S(t-a)}{P(t-a)}E_1(t-a)D_1(a),$$

$$i_j(t, a) = E_j(t-a)D_j(a), j = 2, 3; \text{ and } \quad \forall(t, a) \in \mathbb{R} \times [0, \infty).$$

From the above formulation, we can observe that $i_j(t, a) = i_j(t - a, 0)D_j(a)$, $j = 1, 2, 3$. Moreover, $E_{j+1}(t) = \int_0^\infty \gamma_j(a)D_j(a)E_j(t-a)\mathrm{d}a$, with $j = 1, 2$. Therefore, a straightforward computation gives

$$
\begin{aligned}
E_1(t) = & \int_0^\infty \beta(a)D_1(a)\frac{S(t-a)}{P(t-a)}E_1(t-a)\mathrm{d}a \\
& + \varepsilon\int_0^\infty\int_0^\infty \beta(a)D_2(a)D_1(\eta)\gamma_1(\eta)\frac{S(t-a-\eta)}{P(t-a-\eta)}E_1(t-a-\eta)\mathrm{d}\eta\mathrm{d}a \\
& + \delta\int_0^\infty\int_0^\infty\int_0^\infty \beta(a)D_3(a)D_2(\eta)\gamma_2(\eta)D_1(s)\gamma_1(s)\frac{S(t-a-\eta-s)}{P(t-a-\eta-s)}E_1(t-a-\eta-s)\mathrm{d}s\mathrm{d}\eta\mathrm{d}a.
\end{aligned}
$$
(B.3)

Then, due to Assumption 1.1, (B.3) gives that either E_1 is identically zero or $E_1(t)$ takes on positive values for all $t \in \mathbb{R}$. $\qquad \square$

B.2 Proof of Theorem 2.3 item 1

Let $y_0 \in \mathcal{B}$ and let $\Phi(t, y_0) = (S(t), i_1(t, \cdot), i_2(t, \cdot), i_3(t, \cdot))$ a complete solution in $\mathcal{B}$ passing through y_0 at $t = 0$. By setting $\overline{E}_j = \sup_{t \in \mathbb{R}} E_j(t)$ we successively obtain for all $t \in \mathbb{R}$

$$
\begin{aligned}
E_2(t) &\leq \overline{E}_1 \Gamma_1, \\
E_3(t) &\leq \overline{E}_1 \Gamma_1 \Gamma_2, \\
E_1(t) &\leq \overline{E}_1 \left(\Omega_1 + \varepsilon \Gamma_1 \Omega_2 + \delta \Gamma_1 \Gamma_2 \Omega_3 \right).
\end{aligned}
$$

Taking the supremum on the left-hand side of the E_1 inequality it comes $\overline{E}_1 \leq \overline{E}_1 \mathcal{R}_0$. Then since $\overline{E}_1$ is non-negative and $\mathcal{R}_0 < 1$ it follows $\overline{E}_1 = 0$. Similarly, we also find that $\overline{E}_2 = \overline{E}_3 = 0$. Therefore, the attractor is a compact invariant subset of the space $\mathbb{R} \times \{0\} \times \{0\} \times \{0\}$. Since the only such set is the singleton containing the disease-free equilibrium E^0, this ends the proof of the first part of item 1 of Theorem 2.3.

It remains to prove that E^0 is unstable when $\mathcal{R}_0 > 1$. Notice that the function

$$
\chi(\cdot, E^0) = 1 - [\Omega_1(\cdot) + \varepsilon \Omega_2(\cdot) \Gamma_1(\cdot) + \delta \Omega_3(\cdot) \Gamma_1(\cdot) \Gamma_2(\cdot)]
$$

provided by (B.1) at E^0 is a non-decreasing function on $\mathbb{R}_+$ such that $\chi(0, E^0) = 1 - \mathcal{R}_0 < 0$ and $\lim_{\lambda \to +\infty} \chi(\lambda, E^0) > 0$. As consequence, there exists strictly positive eigenvalue , *i.e.* $\exists \lambda_0 \in \mathbb{R}_+^*$ such that $\chi(\lambda_0, E^0) = 0$. This ends the second part of item 1 of Theorem 2.3.

B.3 Proof of Theorem 2.3 item 2(i) and 2(ii)

Let us start by the proof of item 2(i). By Lemma B.1 the spectrum of the linearized semiflow $\{\Phi(t, \cdot)\}_t$ at the endemic equilibrium E^* is characterized by the following equation, with $\Re(\lambda) > -\mu$:

$$
\begin{aligned}
(\Omega_1(\lambda) + \varepsilon \Omega_2(\lambda) \Gamma_1(\lambda) + \delta \Omega_3(\lambda) \Gamma_1(\lambda) \Gamma_2(\lambda)) &= 1 + \frac{\overline{A}^*}{S^*} + \frac{\overline{A}^* \overline{B}^*}{S^* \left(S^* + \overline{A}^* \right)} \\
+ \frac{\overline{B}^* \left(D_1^\lambda + D_2^\lambda \Gamma_1(\lambda) + D_3^\lambda \Gamma_1(\lambda) \Gamma_2(\lambda) \right)}{S^* + \overline{A}^*}&.
\end{aligned}
\tag{B.4}
$$

Since $\overline{A}^* / S^* = R_0 - 1$, we obtain from the right hand side of (B.4)

$$
|\text{RHS of (B.4)}| > R_0, \quad \forall \lambda \in \mathbb{C} : \Re(\lambda) > -\mu.
\tag{B.5}
$$

Now by contradiction let us assume that there exists $\lambda_0 \in \mathbb{C}$ such that $\Re(\lambda_0) \geq 0$. Then, the left hand side of (B.4) gives

$$
|\text{LHS of (B.4)}| < \Omega_1(0) + \varepsilon \Omega_2(0) \Gamma_1(0) + \delta \Omega_3(0) \Gamma_1(0) \Gamma_2(0) = R_0.
\tag{B.6}
$$

A contradiction holds from (B.5) and (B.6). This ends the proof of Theorem 2.3 item 2(i).

Next, we deal with the proof of Theorem 2.3 item 2(ii). A trivial solution of (1.1) satisfying these initial conditions is given by $i_1(t, \cdot) = i_2(t, \cdot) = i_3(t, \cdot) = 0$ where S is such that $\dot{S} = \Lambda - \mu S$. This trivial solution tends exponentially to E^0. Since solutions to the initial value problem are unique, statement 2(ii) of the theorem follows.

B.4 Proof of Theorem 2.3 item 2(iii)

We assume that the support of at least one of i_{j0}'s has positive measure, and therefore

$$E_1(t) = \int_0^\infty \beta(a) \left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a) \right) \mathrm{d}a$$

takes on positive values for arbitrarily large values of t (by Lem. B.2). Furthermore, Claim A.1 gives that E_1 is Lipschitz, it follows that E_1 is positive on a set of positive measure. In the sequel, when that exists, we set for a given function h: $h^\infty = \limsup_{t\to\infty} h(t)$ and $h_\infty = \liminf_{t\to\infty} h(t)$.

For $\eta_0 > 0$, there exists $t_1 \in \mathbb{R}$ such that $E_1(t) \leq E_1^\infty + \frac{\eta_0}{2}$ for all $t \geq t_1$. Then, it follows from the S-equation of (1.1) that $S_\infty \geq \Lambda \left(\mu + \frac{E_1^\infty}{P_\infty} + \frac{\eta_0}{2} \right)^{-1}$. Thus, there exists $t_2 \geq t_1$ such that

$$S(t) \geq \Lambda \left(\mu + \frac{E_1^\infty}{P_\infty} + \frac{\eta_0}{2} \right)^{-1}, \quad \forall t \geq t_2. \tag{B.7}$$

We perform a time-shift of t_2 on the solution being studied, *i.e.* we replace the initial condition y_0 with $y_1 = \Phi(t_2, y_0)$. The solution passing through y_1 satisfies equations (B.3) and (B.7) for all t, and the bounds E_1^∞ and $P_\infty > 0$ remain valid. Note that (B.3) rewrites

$$
\begin{aligned}
E_1(t) &= \int_0^\infty \beta(a) D_1(a) \frac{S(t-a)}{P(t-a)} E_1(t-a) \mathrm{d}a \\
&+ \varepsilon \int_0^\infty \frac{S(t-\tau)}{P(t-\tau)} E_1(t-\tau) \int_0^\tau \beta(a) D_2(a) D_1(\tau-a) \gamma_1(\tau-a) \mathrm{d}a \mathrm{d}\tau \\
&+ \delta \int_0^\infty \frac{S(t-\tau)}{P(t-\tau)} E_1(t-\tau) \int_0^\tau \int_0^{\tau-a} \beta(a) D_3(a) D_2(\eta) \gamma_2(\eta) D_1(\tau-a-\eta) \gamma_1(\tau-a-\eta) \mathrm{d}\eta \mathrm{d}a \mathrm{d}\tau.
\end{aligned}
$$

From where

$$E_1(t) \geq K \int_0^t u(\tau) E_1(t-\tau) \mathrm{d}\tau, \tag{B.8}$$

wherein $K = \mu \left(\mu + \frac{E_1^\infty}{P_\infty} + \frac{\eta_0}{2} \right)^{-1}$ and

$$u(\tau) = \beta(\tau) D_1(\tau) + \varepsilon \int_0^\tau \beta(a) D_2(a) D_1(\tau-a) \gamma_1(\tau-a) \mathrm{d}a$$

$$+ \delta \int_0^\tau \int_0^{\tau-a} \beta(a) D_3(a) D_2(\eta) \gamma_2(\eta) D_1(\tau-a-\eta) \gamma_1(\tau-a-\eta) \mathrm{d}\eta \mathrm{d}a.$$

Taking the Laplace transform of each side of inequality (B.8) converts the convolution to a product and we obtain for $\lambda \in \mathbb{C}$

$$K \widehat{u}(\lambda) \widehat{E_1}(\lambda) \leq \widehat{E_1}(\lambda).$$

Since E_1 is positive on a set of positive measure then, $\widehat{E_1}$ is strictly positive and the last inequality gives

$$K \int_0^\infty e^{-\lambda\tau} \left(\beta(\tau)D_1(\tau) + \varepsilon \int_0^\tau \beta(a)D_2(a)D_1(\tau-a)\gamma_1(\tau-a)\mathrm{d}a \right.$$
$$\left. + \delta \int_0^\tau \int_0^{\tau-a} \beta(a)D_3(a)D_2(\eta)\gamma_2(\eta)D_1(\tau-a-\eta)\gamma_1(\tau-a-\eta)\mathrm{d}\eta\mathrm{d}a \right) \mathrm{d}\tau \leq 1. \tag{B.9}$$

Changing the order of integration from the left hand side of (B.9), it comes successively

$$\text{LHS of (B.9)} = K \int_0^\infty e^{-\lambda\tau}\beta(\tau)D_1(\tau)\mathrm{d}\tau + K\varepsilon \int_0^\infty \int_0^\infty e^{-\lambda(a+\sigma)}\beta(a)D_2(a)D_1(\sigma)\gamma_1(\sigma)\mathrm{d}\sigma\mathrm{d}a$$
$$+ K\delta \int_0^\infty \int_0^\infty \int_0^\infty e^{-\lambda(\tau+a+\eta)}\beta(a)D_3(a)D_2(\eta)\gamma_2(\eta)D_1(\tau)\gamma_1(\tau)\mathrm{d}\tau\mathrm{d}\eta\mathrm{d}a.$$

From where (B.9) rewrites

$$K \int_0^\infty e^{-\lambda\tau}\beta(\tau)D_1(\tau)\mathrm{d}\tau + K\varepsilon \int_0^\infty \int_0^\infty e^{-\lambda(a+\sigma)}\beta(a)D_2(a)D_1(\sigma)\gamma_1(\sigma)\mathrm{d}\sigma\mathrm{d}a$$
$$+ K\delta \int_0^\infty \int_0^\infty \int_0^\infty e^{-\lambda(\tau+a+\eta)}\beta(a)D_3(a)D_2(\eta)\gamma_2(\eta)D_1(\tau)\gamma_1(\tau)\mathrm{d}\tau\mathrm{d}\eta\mathrm{d}a \leq 1.$$

Taking limits as η_0 and λ tend to zero in the previous inequality, it comes

$$\frac{\mu}{\mu + \frac{E_1^\infty}{P_\infty}} \mathcal{R}_0 \leq 1 \quad i.e. \quad \frac{E_1^\infty}{P_\infty} \geq \mu\left(\mathcal{R}_0 - 1\right).$$

Then, we have the following.

Proposition B.3. *If $\mathcal{R}_0 > 1$, then the semiflow $\{\Phi(t, y_0)\}_t$ is uniformly weakly persistent in the sense that there exists $\nu_0 > 0$ such that*

$$\limsup_{t\to\infty} \int_0^\infty \beta(a)\left(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a)\right)\mathrm{d}a \geq \nu_0.$$

Therefore, the uniform persistence of the semiflow follows from the continuity of E_1 (see Claim A.1), Theorem 2.1 statement 4, Lemma B.2, the uniform weakly persistence of the semiflow (see Prop. B.3) and Theorem 5.2 in [45]. This ends the proof of Theorem 2.3 item 2(iii).

APPENDIX C. BASIC REPRODUCTION NUMBER OF MODEL (1.1)

Let $N(t) = \frac{S(t)}{P(t)}E_1(t)$ be the number of newly HIV infected at time t. Then from (1.1) one has

$$N(t) = E_1(t) = \int_0^\infty \beta(a)(i_1(t,a) + \varepsilon i_2(t,a) + \delta i_3(t,a))\mathrm{d}a$$

where $i_j(t,a)$'s are given by the resolution of the linearized system (1.1) at the disease free equilibrium E^0. Then the Volterra formulation (2.3) yields

$$N(t) = \int_0^t \beta(a)\left(D_1(a)E_1(t-a) + \varepsilon D_2(a)E_2(t-a) + \delta D_3(a)E_3(t-a)\right)\mathrm{d}a + N_0(t), \tag{C.1}$$

with $N_0(t) = \int_t^\infty \beta(a) \left(i_{10}(a-t)\frac{D_1(a)}{D_1(a-t)} + \varepsilon i_{20}(a-t)\frac{D_2(a)}{D_2(a-t)} + \delta i_{30}(a-t)\frac{D_3(a)}{D_3(a-t)} \right) \mathrm{d}a$.

Further, we have

$$E_2(t) = \int_0^t \gamma_1(a)D_1(a)E_1(t-a)\mathrm{d}a + H_2(t),$$

$$E_3(t) = \int_0^t \gamma_2(a)D_2(a)E_2(t-a)\mathrm{d}a + \int_t^\infty \gamma_2(a)i_{20}(a-t)\frac{D_2(a)}{D_2(a-t)}\mathrm{d}a$$

$$= \int_0^t E_1(t-a)U(a)\mathrm{d}a + H_3(t),$$

with $H_2(t) = \int_t^\infty \gamma_1(a)i_{10}(a-t)\frac{D_1(a)}{D_1(a-t)}\mathrm{d}a$, $H_3(t) = \int_0^t \gamma_2(a)D_2(a)H_2(t-a)\mathrm{d}a + \int_t^\infty \gamma_2(a)i_{20}(a-t)\frac{D_2(a)}{D_2(a-t)}\mathrm{d}a$, and $U(a) = \int_0^a \gamma_1(a-\sigma)D_1(a-\sigma)\gamma_2(\sigma)D_2(\sigma)\mathrm{d}\sigma$.

Therefore,

$$\int_0^t \beta(a)D_2(a)E_2(t-a)\mathrm{d}a = \int_0^t E_1(t-a)\int_0^a D_1(a-\sigma)\gamma_1(a-\sigma)\beta(\sigma)D_2(\sigma)\mathrm{d}\sigma\mathrm{d}a$$

$$+ \int_0^t \beta(a)D_2(a)H_2(t-a)\mathrm{d}a,$$

and

$$\int_0^t \beta(a)D_3(a)E_3(t-a)\mathrm{d}a = \int_0^t E_1(t-a)\int_0^a U(a-\sigma)\beta(\sigma)D_3(\sigma)\mathrm{d}\sigma\mathrm{d}a$$

$$+ \int_0^t \beta(a)D_3(a)H_3(t-a)\mathrm{d}a.$$

Consequently, (C.1) rewrites

$$N(t) = \int_0^t K(a)E_1(t-a)\mathrm{d}a + N_0(t)$$

$$+ \varepsilon \int_0^t \beta(a)D_2(a)H_2(t-a)\mathrm{d}a + \delta \int_0^t \beta(a)D_3(a)H_3(t-a)\mathrm{d}a,$$

wherein

$$K(a) = \beta(a)D_1(a) + \varepsilon \int_0^a D_1(a-\sigma)\gamma_1(a-\sigma)\beta(\sigma)D_2(\sigma)\mathrm{d}\sigma\mathrm{d}a + \delta \int_0^a U(a-\sigma)\beta(\sigma)D_3(\sigma)\mathrm{d}\sigma.$$

From the above formulation, the basic reproduction number is calculated as

$$
\begin{aligned}
R_0 &= \int_0^\infty K(a)\mathrm{d}a \\
&= \int_0^\infty \beta(a)D_1(a)\mathrm{d}a + \varepsilon \int_0^\infty \int_0^a D_1(a-\sigma)\gamma_1(a-\sigma)\beta(\sigma)D_2(\sigma)\mathrm{d}\sigma\mathrm{d}a \\
&\quad + \delta \int_0^\infty \int_0^a \beta(\sigma)D_3(\sigma)U(a-\sigma)\mathrm{d}\sigma\mathrm{d}a \\
&= \int_0^\infty \beta(a)D_1(a)\mathrm{d}a + \varepsilon \int_0^\infty \beta(a)D_2(a) \int_a^\infty D_1(\sigma-a)\gamma_1(\sigma-a)\mathrm{d}\sigma\mathrm{d}a \\
&\quad + \delta \int_0^\infty \beta(a)D_3(a) \int_a^\infty U(\sigma-a)\mathrm{d}\sigma\mathrm{d}a \\
&= \int_0^\infty \beta(a)D_1(a)\mathrm{d}a + \varepsilon \int_0^\infty \beta(a)D_2(a)\mathrm{d}a \int_0^\infty D_1(\sigma)\gamma_1(\sigma)\mathrm{d}\sigma \\
&\quad + \delta \int_0^\infty \beta(a)D_3(a)\mathrm{d}a \int_0^\infty \gamma_2(\sigma)D_2(\sigma)\mathrm{d}\sigma \int_0^\infty \gamma_1(\tau)D_1(\tau)\mathrm{d}\tau, \\
&= \Omega_1 + \varepsilon\Gamma_1\Omega_2 + \delta\Gamma_1\Gamma_2\Omega_3.
\end{aligned}
$$

APPENDIX D. EXISTENCE OF AN OPTIMAL CONTROL

Let $\mathcal{V} = \{1, 1.TF, 2, 2.TF, 3, 1.TS, 2.TS\}$ be the set of indexes. Let us consider a control $h \in \mathcal{U}$ and denote by $w^h(t,a) = \left(S^h(t), (i_j^h(t,a))_{j\in\mathcal{V}}\right)$, (resp. $\lambda^h = \left(\lambda_S^h(t), (\lambda_j^h(t,a))_{j\in\mathcal{V}}\right)$), the corresponding state, (resp. adjoint), vector variable given by (4.3) and (4.6)–(4.7). Let us define the map $\mathcal{L} : L^1((0,T_f), \mathbb{R}^3) \longrightarrow L^\infty((0,T_f), \mathbb{R}^3)$ by $\mathcal{L}(u_1, u_2, u_3) = (\mathcal{L}_1 u_1, \mathcal{L}_2 u_2, \mathcal{L}_{2.TF} u_3)$, where

$$
\mathcal{L}_j u = \begin{cases} 0, & \text{if} \quad u < 0, \\ u, & \text{if} \quad 0 \le u < h_j^{\max}, \\ h_j^{\max}, & \text{if} \quad u \ge h_j^{\max}, \end{cases} \quad j \in \{1; 2; 2.TF\}.
$$

By setting $\mathcal{X} := (0, T_f) \times \left(Q_{T_f}\right)^7$ with $Q_{T_f} = (0, T_f) \times (0, \infty)$, we define the norms $\|\cdot\|_{L^1(\mathcal{X})}$ and $\|\cdot\|_{L^\infty(\mathcal{X})}$ such that for a given vector function $(y, x) := (y, (x_j)_{j=1,\dots,7})$,

$$
\|(y, x)\|_{L^1(\mathcal{X})} = \int_0^{T_f} |y(t)|\mathrm{d}t + \sum_{j=1}^7 \int_0^{T_f} \int_0^\infty |x_j(t, a)|\mathrm{d}t\mathrm{d}a,
$$

$$
\|(y, x)\|_{L^\infty(\mathcal{X})} = \sup_{t \in [0, T_f]} |y(t)| + \sum_{j=1}^7 \sup_{t \in [0, T_f]} \int_0^\infty |x_j(t, a)|\mathrm{d}a.
$$

In the same way, define the norms $\|.\|_{L^1(Q_{T_f})}$ and $\|.\|_{L^\infty(Q_{T_f})}$. We embed our optimal problem in the space $L^1(0, T_f)$ by defining the function

$$
\mathcal{J}(h) = \begin{cases} J(h), & \text{if} \ h \in \mathcal{U}, \\ +\infty, & \text{if} \ h \notin \mathcal{U}. \end{cases}
$$

To prove the existence of the optimal control, let us introduce the first preliminary result.

Lemma D.1. *Let T_f be sufficiently small.*
1. The map $h \in \mathcal{U} \to w^h \in L^1(\mathcal{X})$ is Lipschitz for the norms $\|\cdot\|_{L^1}$ and $\|\cdot\|_{L^\infty}$ in the following way:

$$\|w^h - w^v\|_{L^1(\mathcal{X})} \leq T_f C \|h - v\|_{L^1(0,T_f)} \ and \ \|w^h - w^f\|_{L^\infty(\mathcal{X})} \leq T_f C \|h - f\|_{L^\infty(0,T_f)}$$

for all $h, f \in \mathcal{U}$ and where $C > 0$ is a constant.
2. For $h \in \mathcal{U}$, the adjoint system (4.6)–(4.7) has a weak solution λ^h in $L^\infty(\mathcal{X})$ such that

$$||\lambda^h - \lambda^f||_{L^\infty(\mathcal{X})} \leq T_f C ||h - f||_{L^\infty(0,T_f)},$$

for all $h, f \in \mathcal{U}$ and $C > 0$.
3. The functional $\mathcal{J}(h)$ is lower semicontinuous with respect to $L^1(0, T_f)$ convergence.

Proof. Recalling $E_1(t) = \int_0^\infty \beta(a) \left[i_1 + i_{1,TF} + \varepsilon(i_2 + i_{2,TF}) + \delta i_3 \right](t, a) \mathrm{d}a$, and since the total population P is bounded by positive constants, we easily find a constant $C_0 > 0$ such that

$$\left| \frac{E_1^h(t)}{P^h(t)} - \frac{E_1^f(t)}{P^f(t)} \right| \leq C_0 \|\beta\|_\infty \sum_{v \in \{i_1, i_{1,TF}, i_2, i_{2,TF}, i_3\}} \|v^h(t, \cdot) - v^f(t, \cdot)\|_{L^1(0,\infty)}.$$

Moreover, by using the same argument as in item 2. of Theorem 2.1, we can find a positive constant M_0 such that $|S^h(t)| + \sum_{v \in \{i_1, i_{1,TF}, i_2, i_{2,TF}, i_3\}} \|v^h(t, \cdot)\|_{L^1(0,\infty)} \leq M_0$ for all $h \in \mathcal{U}$ and a.e. t.

Then, integrating the S-component of (4.1) it comes

$$\begin{aligned}
|S^h(t) - S^f(t)| &\leq \|\beta\|_\infty C_0 M_0 \int_0^t e^{-\mu\tau} |S^h(\tau) - S^f(\tau)| \mathrm{d}\tau \\
&+ \|\beta\|_\infty C_0 M_0 \int_0^t e^{-\mu\tau} \sum_{v \in \{i_1, i_{1,TF}, i_2, i_{2,TF}, i_3\}} \|v^h(\tau, \cdot) - v^f(\tau, \cdot)\|_{L^1(0,\infty)} \mathrm{d}\tau.
\end{aligned} \tag{D.1}$$

Moreover, the Volterra integral formulation of system (4.3) gives for i_1-equation

$$i_1^h(t, a) = \begin{cases} i_{10}(a - t) \frac{D_1(a)}{D_1(a-t)}, & \text{for } t < a, \\ (1 - h_1(t - a)) \frac{S^h(t-a)}{P^h(t-a)} E_1^h(t - a) D_1(a), & \text{for } t \geq a \end{cases}$$

From where,

$$\begin{aligned}
\|i_1^h(t, \cdot) - i_1^f(t, \cdot)\|_{L^1(0,\infty)} &\leq M_0^2 \|\beta\|_\infty \|h_1 - f_1\|_{L^1(0,T_f)} \\
&+ C_0 M_0 T_f \|\beta\|_\infty \left(\sum_{v \in \{i_1, i_{1,TF}, i_2, i_{2,TF}, i_3\}} \|v^h(t, \cdot) - v^f(t, \cdot)\|_{L^1(0,\infty)} + |S^h(t) - S^f(t)| \right).
\end{aligned} \tag{D.2}$$

Therefore, applying same argument for $i_{1,TF}$, i_2, $i_{2,TF}$ and i_3 as for estimates $(D.2)$ and combining with $(D.1)$, it follows that for T_f sufficiently small,

$$||w^h - w^v||_{L^1(\mathcal{X})} \leq T_f C ||h - v||_{L^1(0,T_f)}.$$

The same argument is then applied for the norm L^∞ and for item 2. It remains to prove item 3.

We suppose that $h_n := (h_{1n}, h_{2n}, h_{2n,TF}) \to h := (h_1, h_2, h_{2,TF})$ in $L^1(0, T_f)$. Possibly along a subsequence and using the same notation, $h_n^2 \to h^2$ a.e. on $(0, T_f)$ by (see [15], p. 21). By Lebesgue's dominated convergence

theorem, it comes $\lim_{n\to\infty}||h_n^2||_{L^1(0,T_f)} = ||h^2||_{L^1(0,T_f)}$. We have the similar argument for $||f^2||_{L^1(0,T_f)}$. These handle the convergence of the squared terms in the functional.

Next, we illustrate the convergence of one term in the functional,

$$\begin{aligned}||B\gamma_2(i_2^{h_n} - i_2^h)||_{L^1(Q_{T_f})} &\leq ||B||_\infty||\gamma_2||_\infty||w^{h_n} - w^h||_{L^1(\mathcal{X})}\\ &\leq CT_f||h_n - h||_{L^1(0,T_f)}.\end{aligned}$$

Therefore,

$$|\mathcal{J}(h_n) - \mathcal{J}(h)| \leq CT_f||h_n - h||_{L^1(0,T_f)}.$$

From where we deduce the lower semi-continuity, $\mathcal{J}(h) \leq \liminf_{n\to\infty} \mathcal{J}(h_n).$ □

The functional $\mathcal{J} : (0, T_f) \to (-\infty, \infty]$ is lower semi-continuous with respect to strong L^1 convergence but not with respect to weak L^1 convergence. Thus, in general it does not attain its infimum on $(0, T_f)$. Thus, we circumvent this situation by using the Ekeland's variational principle (see [14]): for $\delta > 0$, there exists h_δ in $L^1(0, T_f)$ such that

$$\mathcal{J}(h_\delta) \leq \inf_{h\in\mathcal{U}} \mathcal{J}(h) + \delta, \tag{D.3}$$

$$\mathcal{J}(h_\delta) = \min_{h\in\mathcal{U}}\left\{\mathcal{J}(h) + \sqrt{\delta}||h_\delta - h||_{L^1(0,T_f)}\right\}. \tag{D.4}$$

Note that, by (D.4), the perturbed functional

$$\mathcal{J}_\delta(h) = \mathcal{J}(h) + \sqrt{\delta}||h_\delta - h||_{L^1(0,T_f)}$$

attains its infimum at h_δ. By the same argument as in Section 4.3, and using the projection map $\mathcal{L}$ on $\mathcal{U}$, it comes that

Lemma D.2. *If h_δ is an optimal control minimizing the functional $\mathcal{J}_\delta(h)$, then*

$$h_\delta = \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_1^\delta}{2C_1}, \hat{h}_2(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_2^\delta}{2C_2}, \hat{h}_{2,TF}(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_3^\delta}{2C_3}\right);$$

where $\pi_j^\delta \in L^\infty(0, T_f)$, with $|\pi_j^\delta(\cdot)| \leq 1$ and

$$\hat{h}_1(\lambda^{h_\delta}) = \frac{\left(\lambda_{i_1}(t,0) - p_1\lambda_{i_1,TF}(t,0) - (1-p_1)\lambda_{i_1,TS}(t,0)\right)}{2C_1}\frac{SE_1^{h_\delta}}{P^{h_\delta}},$$

$$\hat{h}_2(\lambda^{h_\delta}) = \frac{\left(\lambda_{i_2}(t,0) - p_2\lambda_{i_2,TF}(t,0) - (1-p_2)\lambda_{i_2,TS}(t,0)\right)}{2C_2}E_2^{h_\delta},$$

$$\hat{h}_{2,TF}(\lambda^{h_\delta}) = \frac{\left(\lambda_{i_2}(t,0) - p_{2,TF}\lambda_{i_2,TF}(t,0) - (1-p_{2,TF})\lambda_{i_2,TS}(t,0)\right)}{2C_3}E_{2,TF}^{h_\delta}.$$

By Lemmas D.1 and D.2, we are now ready to prove the existence and uniqueness of an optimal controller (namely Thm. 4.3).

Poof of Theorem 4.3. We start with the uniqueness by defining $\mathcal{F} : \mathcal{U} \longrightarrow \mathcal{U}$ by

$$\mathcal{F}(h) = \mathcal{L}\left(\hat{h}_1(\lambda^h), \hat{h}_2(\lambda^h), \hat{h}_{2,TF}(\lambda^h)\right)$$

wherein w^h and λ^h are state and adjoint solutions corresponding to h as in previous sections. Using the Lipschitz properties of w^h and λ^h (see Lem. D.1), for $h, \bar{h} \in \mathcal{U}$, we find $C > 0$ such that for T_f sufficiently small

$$
\begin{aligned}
\|\mathcal{L}_1(h_1) - \mathcal{L}_1(\bar{h}_1)\|_{L^\infty(0,T_f)} &\leq \frac{C_0 \max\left(M_0, M_0^2\right)}{2C_1} \\
&\quad \times \left(\|w^h - w^{\bar{h}}\|_{L^\infty(Q_{T_f})} + \|\lambda^h - \lambda^{\bar{h}}\|_{L^\infty(Q_{T_f})} \right) \\
&\leq \frac{CT_f}{2C_1} \|h - \bar{h}\|_{L^\infty(0,T_f)}.
\end{aligned}
$$

Applying the same argument for $\mathcal{L}_2$ and $\mathcal{L}_{2,TF}$, it comes

$$
\|\mathcal{F}(h) - \mathcal{F}(\bar{h})\|_{L^\infty(0,T_f)} \leq CT_f \|h - \bar{h}\|_{L^\infty(0,T_f)} \left(\frac{1}{2C_1} + \frac{1}{2C_2} + \frac{1}{2C_3} \right) \tag{D.5}
$$

where the positive constant C depends on the L^∞ bounds on the state and the adjoint solutions and Lipschitz constants. Therefore, if $T_f < \frac{2}{C}(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3})^{-1}$, thus the map $\mathcal{F}$ has a unique fixed point h^*.

It remains to prove that this fixed point is an optimal controller. For that end, we used the approximate minimizers h_δ from Ekeland's variational principle. From Lemma D.2 and the contraction property of $\mathcal{F}$, we have

$$
\begin{aligned}
&\left\| \mathcal{F}(h_\delta) - \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_1^\delta}{2C_1}, \hat{h}_2(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_2^\delta}{2C_2}, \hat{h}_{2,TF}(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_3^\delta}{2C_3} \right) \right\|_{L^\infty(0,T_f)} = \\
&\left\| \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}), \hat{h}_2(\lambda^{h_\delta}), \hat{h}_{2,TF}(\lambda^{h_\delta}) \right) \right. \\
&\left. - \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_1^\delta}{2C_1}, \hat{h}_2(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_2^\delta}{2C_2}, \hat{h}_{2,TF}(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_3^\delta}{2C_3} \right) \right\|_{L^\infty(0,T_f)} \\
&\leq \sum_{j=1}^3 \left\| \frac{\sqrt{\delta}\pi_j^\delta}{2C_j} \right\|_{L^\infty(0,T_f)} \leq \sqrt{\delta} \left(\frac{1}{2C_1} + \frac{1}{2C_2} + \frac{1}{2C_3} \right).
\end{aligned} \tag{D.6}
$$

Consequently, from (D.5) and (D.6), we have

$$
\begin{aligned}
&\|h^* - h_\delta\|_{L^\infty(0,T_f)} = \\
&\left\| \mathcal{F}(h^*) - \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_1^\delta}{2C_1}, \hat{h}_2(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_2^\delta}{2C_2}, \hat{h}_{2,TF}(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_3^\delta}{2C_3} \right) \right\|_{L^\infty(0,T_f)} \\
&\leq \|\mathcal{F}(h^*) - \mathcal{F}(h_\delta)\|_{L^\infty(0,T_f)} \\
&+ \left\| \mathcal{F}(h_\delta) - \mathcal{L}\left(\hat{h}_1(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_1^\delta}{2C_1}, \hat{h}_2(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_2^\delta}{2C_2}, \hat{h}_{2,TF}(\lambda^{h_\delta}) + \frac{\sqrt{\delta}\pi_3^\delta}{2C_3} \right) \right\|_{L^\infty(0,T_f)} \\
&\leq \left(CT_f \|h^* - h_\delta\|_{L^\infty(0,T_f)} + \sqrt{\delta} \right) \left(\frac{1}{2C_1} + \frac{1}{2C_2} + \frac{1}{2C_3} \right).
\end{aligned}
$$

Since $CT_f\left(\frac{1}{2C_1} + \frac{1}{2C_2} + \frac{1}{2C_3} \right)$ is sufficiently small, it comes

$$
\|h^* - h_\delta\|_{L^\infty(0,T_f)} \leq \sqrt{\delta} \left[1 - CT_f \sum_{j=1}^3 \frac{1}{2C_j} \right]^{-1} \sum_{j=1}^3 \frac{1}{2C_j},
$$

which gives $h_\delta \longrightarrow h^*$ in $L^\infty(0,T_f)$ (as $\delta \longrightarrow 0$) and by (D.3)

$$
\mathcal{J}(h^*) = \inf_{h \in \mathcal{U}} \mathcal{J}(h).
$$

$\square$

REFERENCES

[1] Aide Suisse Contre le SIDA. (2018). Evolution d'une infection par le VIH. www.aids.ch/fr/vivre-avec-vih/aspects-medicaux/evolution.php.

[2] AIDS info. (2018). Offering Information on HIV/AIDS Treatment, Prevention and Research. aidsinfo.nih.gov/understanding-hiv-aids/fact-sheets/21/51/hiv-treatment--the-basics.

[3] S. Anita, Vol. 11 of *Analysis and control of age-dependent population dynamics*. Springer Science & Business Media (2000).

[4] V. Barbu and M. Iannelli, Optimal control of population dynamics. *J. Optim. Theory Appl.* **102** (1999) 1–14.

[5] T. Bernard, K. Diop and P. Vinard, The cost of universal free access for treating HIV/AIDS in low-income countries: the case of Senegal (2008). http://hal.ird.fr/ird-00403656.

[6] Centers for Disease Control and Prevention (CDC). (2018). HIV Prevention. https://www.cdc.gov/hiv/basivs/prevention.html.

[7] J. Chu, A. Ducrot, P. Magal and S. Ruan, Hopf bifurcation in a size-structured population dynamic model with random growth. *J. Differ. Equ.* **247** (2009) 956–1000.

[8] G. Da Prato and M. Iannelli, Boundary control problems for age-dependent equations. In Vol. 155 of *Evolution equations, control theory and biomathematics*. Marcel Dekker (1993) 91–100.

[9] C.A. da Silva Filho and J.L. Boldrini, An analysis of an optimal control problem for mosquito populations. *Nonlinear Anal.: Real World Appl.* **42** (2018) 353–377.

[10] R. Djidjou-Demasse, J.J. Tewa, S. Bowong and Y. Emvudu, Optimal control for an age-structured model for the transmission of hepatitis B. *J. Math. Biol.* **73** (2016) 305–333.

[11] O. Diekmann, J.A.P. Heesterbeek and J.A. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28** (1990) 365–382.

[12] A. Ducrot, Z. Liu and P. Magal, Essential growth rate for bounded linear perturbation of non-densely defined Cauchy problems. *J. Math. Anal. Appl.* **341** (2008) 501–518.

[13] J.W. Eaton and T.B. Hallett, Why the proportion of transmission during early-stage HIV infection does not predict the long-term impact of treatment on HIV incidence. *Proc. Natl. Acad. Sci.* **111** (2014) 16202–16207.

[14] I. Ekeland, On the variational principle. *J. Math. Anal. Appl.* **47** (1974) 324–353.

[15] L.C. Evans and R.F. Gariepy, Measure Theory and Fini Properties of Functions. CRC Press, Boca Raton (1992).

[16] G. Feichtinger, G. Tragler and V.M. Veliov, Optimality conditions for age-structured control systems. *J. Math. Anal. Appl.* **288** (2003) 47–68.

[17] K.R. Fister and S. Lenhart, Optimal control of a competitive system with age-structure. *J. Math. Anal. Appl.* **291** (2004) 526–537.

[18] A.B. Gumel, C.C. McCluskey and P. van den Driessche, Mathematical study of a staged-progression HIV model with imperfect vaccine. *Bull. Math. Biol.* **68** (2006) 2105–2128.

[19] H. Guo and M.Y. Li, Global dynamics of a staged-progression model for HIV/AIDS with amelioration. *Nonlinear Anal.: Real World Appl.* **12** (2011) 2529–2540.

[20] J.K. Hale, J.P. LaSalle and M. Slemrod, Theory of a general class of dissipative processes. *J. Math. Anal. Appl.* **39** (1972) 177–191.

[21] T.D. Hollingsworth, R.M. Anderson and C. Fraser, HIV-1 transmission, by stage of infection. *J. Infect. Dis.* **198** (2008) 687–693.

[22] J.F. Hutchinson, The biology and evolution of HIV. *Annu. Rev. Anthropol.* **30** (2001) 85–108.

[23] J.M. Hyman, J. Li and E.A. Stanley, The differential infectivity and staged progression models for the transmission of HIV12. *Math. Biosci.* **155** (1999) 77–109.

[24] J.M. Hyman, J. Li and E.A. Stanley, The differential infectivity and staged progression models for the transmission of HIV. *Math. Biosci.* **155** (1999) 77–109.

[25] M. Iannelli, Mathematical Theory of Age-Structured Population Dynamics. Giadini Editori e Stampatori, Pisa (1994).

[26] H. Inaba, On a new perspective of the basic reproduction number in heterogeneous environments. *J. Math. Biol.* **65** (2012) 309–348.

[27] M.E. Kretzschmar, M.F.S. van der Loeff, P.J. Birrell, D. De Angelis and R.A. Coutinho, Prospects of elimination of HIV with test-and-treat strategy. *Proc. Natl. Acad. Sci.* **110** (2013) 15538–15543.

[28] O.A. Ladyzhenskaya, On the determination of minimal global attractors for the Navier-Stokes and other partial differential equations. *Uspekhi Mat. Nauk* **42** (1987) 25–60.

[29] S. Lenhart and J.T. Workman, Optimal control applied to biological models. CRC Press (2007).

[30] X. Lin, H.W. Hethcote and P. Van den Driessche, An epidemiological model for HIV/AIDS with proportional recruitment. *Math. Biosci.* **118** (1993) 181–195.

[31] P. Magal and S. Ruan, On semilinear Cauchy problems with non-dense domain. *Adv. Differ. Equ.* **14** (2009) 1041–1084.

[32] P. Magal and S. Ruan, Center manifolds for semilinear equations with non-dense domain and applications to Hopf bifurcation in age structured models. *Am. Math. Soc.* (2009).

[33] P. Magal, C.C. McCluskey and G.F. Webb, Lyapunov functional and global asymptotic stability for an infection-age model. *Appl. Anal.* **89** (2010) 1109–1140.

[34] P. Magal, Compact attractors for time-periodic age-structured population models. *Electr. J. Differ. Equ.* (2001).

[35] C.C. McCluskey, A model of HIV/AIDS with staged progression and amelioration. *Math. Biosci.* **181** (2003) 1–16.

[36] E. Numfor, S. Bhattacharya, M. Martcheva and S. Lenhart, Optimal control in multi-group coupled within-host and between-host models. *Electr. J. Differ. Equ.* **2016** (2016) 87–117.

[37] R.N. Nsubuga, R.G. White, B.N. Mayanja and L.A. Shafer, Estimation of the HIV basic reproduction number in rural South West Uganda: 1991–2008. *PloS one* **9** (2014) e83778.

[38] A. Pazy, Semigroups of Linear Operator and Applications to Partial Differential Equations. *Appl. Math. Sci.* **44** (1983).

[39] A.S. Perelson and P.W. Nelson, Mathematical analysis of HIV-1 dynamics in vivo. *SIAM Rev.* **41** (1999) 3–44.

[40] S.D. Pinkerton, Probability of HIV transmission during acute infection in Rakai. *Natl. Inst. Health Uganda. AIDS Behavior* **12** (2008) 677–684.

[41] G. Rozhnova, M.F.S. van der Loeff, J.C. Heijne and M.E. Kretzschmar, Impact of heterogeneity in sexual behavior on effectiveness in reducing HIV transmission with test-and-treat strategy. *PLoS Comput. Biol.* **12** (2016) e1005012.

[42] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, *et al.*, Global sensitivity analysis: the primer. John Wiley & Sons (2008).

[43] G.R. Sell and Y. You, Dynamics of Evolutionary Equations. Springer, New York (2002).

[44] M. Shen, Y. Xiao and L. Rong, Global stability of an infection-age structured HIV-1 model linking within-host and between-host dynamics. *Math. Biosci.* **263** (2015) 37–50.

[45] H.L. Smith and H.R. Thieme, Vol. 118 of Dynamical systems and population persistence. American Mathematical Soc. (2011).

[46] H.R. Thieme, "Integrated semigroups" and integrated solutions to abstract Cauchy problems. *J. Math. Anal. Appl.* **152** (1990) 416–447.

[47] H.R. Thieme, Quasi-compact semigroups via bounded perturbation. Advances in Mathematical Population Dynamics-Molecules, Cells and Man, edited by O. Arino, D. Axelrod, and M. Kimmel. *Worlds Scientific* (1997) 691–713.

[48] H.R. Thieme, Global stability of the endemic equilibrium in infinite dimension: Lyapunov functions and positive operators. *J. Differ. Equ.* **250** (2011) 3772–3801.

[49] United Nations Programme on HIV/AIDS (UNAIDS) (2016). Global AIDS Update 2016. www.unaids.org.

[50] G.F. Webb, Theory of nonlinear age-dependent population dynamics. CRC Press (1985).

[51] R. Weston and B. Marett, HIV infection pathology and disease progression. Vol. 1 of Clinical Pharmacist (2009).

[52] World Health Organization (WHO) (2017). www.who.int/features/qa/71/fr/.