

# The impact of self-protective measures in the optimal interventions for controlling infectious diseases of human population

Semu Mitiku Kassa · Aziz Ouhinou

Received: 15 March 2013 / Revised: 22 January 2014 / Published online: 11 February 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** A mathematical model for infectious disease epidemics with behaviour change and treatment is formulated and analysed. It is indicated that behaviour modification by the population has a significant impact on the dynamics of the disease. Moreover, an optimal control theory is applied to propose the best possible combination of efforts in controlling a disease. It is shown that it may not be necessary to continuously apply treatment at a full rate to eradicate the disease, if the effort is supported by effective behaviour modification strategies.

**Keywords** Dynamical systems · Epidemiological models · Infectious diseases · Behaviour change · Optimal control

**Mathematics Subject Classification** 34A34 · 37G10 · 92B05 · 93A30 · 93B40 · 93C15

## 1 Introduction

In the last two decades epidemiological models have made important advances by identifying the role of heterogeneous contact processes in the spread of infectious diseases. Using such quantitative models researchers have come up with recommendations of various intervention mechanisms in controlling different types of diseases (Gaff and Schaefer 2009; Hethcote 2000; Brauer et al. 2008). Nonetheless, many

---

S. M. Kassa (✉)

Department of Mathematics, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia  
e-mail: semu.mitiku@aau.edu.et

A. Ouhinou

Department of Mathematics, Faculty of Science and technology,  
University of Sultane Moulay Slimane, B.P. 523, Beni-Mellal, Morocco

existing models are based on the assumption that the behaviour of individuals to protect themselves against an infectious disease remains constant or unchanged in the course of the outbreak.

In practice, however, the change in behaviour of human population has influenced on the spread of various infectious diseases. The effect of human learning behaviour and adoption of self-protective measures against an epidemic have been witnessed during the outbreak of influenza A (H1N1) in 2009 (cf. [Funk et al. 2010](#) and the references therein). Many people have been seen wearing face-masks and have changed their travelling behaviour during the 2003 SARS epidemics ([Lau et al. 2005](#)). Such endogenous self-protective measures have had a noticeable impact on the spread of the disease. Individuals are most likely to change their risky behaviour when the morbidity of the disease or the perception of risk is very high. For the effects of behaviour change to increase, populations need to receive more accurate information about the disease ([Chen 2009](#)), and the various self protective mechanisms that has to be adopted should also be at a reasonable level of efficacy. Modelling and including the effect of behavioural changes during an epidemic can give the dynamics very different form, and closer to the reality than the predictions made using models with unchanged behaviour ([Ferguson 2007](#); [Kassa and Ouhinou 2011](#)).

Some individuals in the population start taking self-initiated actions to try to reduce their risk of contracting a disease once they have a first hand experience of the disease, i.e., if one of their family members or one of their loved ones is affected because of the disease. Others may change their risky behaviour only when the risk is very high and may need the effort of persuasive agents (like government agencies, or public health organizations). But the ratio of the population to either side or the rate of the behaviour change may vary from disease to disease depending on its lethality. Using individual contact tracing network model [Funk et al. \(Funk et al. 2009\)](#) concluded that awareness of a disease can possibly reduce the susceptibility of the population and in some special cases can also stop the spread of the disease. Using utility analysis techniques, [Chen \(2004, 2006\)](#) showed that there is a unique endemic equilibrium prevalence when the basic reproductive number of a sexually transmitted disease is strictly greater than unity. Moreover, [Chen](#) has also concluded in [Chen \(2009\)](#) that the likelihood of eradicating an infectious disease through behavioural changes depends critically on the amount and quality of information individuals have access to. These considerations suggest that including self protective measures in the model will have a large impact in the study of the dynamics of infectious diseases. However, awareness about a disease alone does not directly lead to using self-protective measures. As researchers in health behaviour and health education indicate, the majority of the population need persuasive interventions to decide to change their behaviour before it is too late ([Bertrand 2004](#); [Oldenburg and Glanz 2008](#)).

Evidences show that behaviour change is prevalence-elastic. That means, more individuals adopt self-protective measures as prevalence of the disease increases in their area ([Ahituv et al. 1996](#); [Philipson 1996](#)). However, most mathematical models do not take the effect of these behavioural changes in to account.

To react against an infectious disease and eventually take protective action, one needs to clearly experience danger or get concrete information about its happening. In this regard, the qualitative nature of information about a disease and the possi-

ble protective mechanisms are very important to investigate. An individual can get information about the disease either through word of mouth from people that have infection experience, or through mass campaigns that are targeting the population to inform about the danger of the disease. The former mechanism is dependent mainly on social networking and the social culture of the society (Funk et al. 2009), while the latter depends on the quality of the campaign or public effort to bring more impact, and expenditure in dissemination of the information. The preventive interventions in curbing disease epidemics requires a huge effort and investment from public health sectors in preventing the disease. All these require some public fund expenditure. To allocate the meagre public resource in protective mechanisms one need to know the optimum way to do it, especially when it is combined with pharmaceutical interventions. On the other hand informing the population about the outbreak of a certain disease may only put them in panic as has been seen in Surat, India when people got information about a presumed outbreak of bubonic plague in the year 1994 (Campbell and Hughes 1995). But the various ways of protective mechanisms in protecting oneself from the disease epidemics should also be devised and disseminated into the population. The more choices of alternative measures for behaviour modification offered to the population the more people decide to use at least one of them and hence more effective than rigidly prescribing a single behaviour for change (see Turner et al. 1989, pp. 276–279). Therefore, working in enhancing the various protective measures and giving alternatives that best protect the population is very important; as well as alternatives that can be better adhered to for longer duration by the susceptible individuals in the population. It is intuitively clear that the more efficacious, on average, the protective measures are, the more people decide to use one of them and the more people adhere to it for longer.

Joshi et al. (2008) has formulated a mathematical model to trace the effect of information campaigns on the dynamics of HIV/AIDS in Uganda. In the model the authors assume that people take self-protective actions only when they interact with the educational campaign class. Moreover, the model does not incorporate the efficacy level of the protective mechanisms, which is useful to be considered by policy makers to take any action. Fenichel et al. (2011) also considered a model with varying contact rate between individuals depending on their health status to address the effect of change in human behaviour in disease dynamics. However, the contact rate affects only the number of contacts between individuals but does not take in to account the effect of using protective devices by individuals while contacting others.

To evaluate the effects of the overall public health measures and to plan effective pharmaceutical interventions to control the epidemic of an infectious disease, it is very important to explore the contribution of behaviour modifications and self-initiated actions against the disease. In Kassa and Ouhinou (2011) an epidemiological model with the inclusion of effects of self-protective measures has been formulated. However, the model does not include the effect of treatment to the infected once. The motivation for the model in the above mentioned paper was to see the sensitivity of the dynamics to either of the rate of information diffusion or the average strength of the protective mechanisms; while here we are interested in optimally combining self-protective measures with pharmaceutical interventions. Thus, in this paper we extend our previous result in Kassa and Ouhinou (2011) to incorporate treatment or any other intervention on the infected population. Using this model we investigate the dynam-

ics of an epidemic and also apply optimal control analysis to propose an efficient combined public health intervention strategy to control the spread of the disease.

The paper is organized as follows, in Sect. 2, we describe and formulate the mathematical model consisting of a system of ordinary differential equations that describe the impact of behaviour change modification when treatment is also considered as a possible intervention mechanism. The mathematical analysis of the model is discussed in Sect. 3. Section 4 is devoted to the optimal control of the model and to the simulation results of the optimality system. We conclude the paper with a discussion in Sect. 5.

## 2 The mathematical model with change behaviour and treatment

In this section we formulate and analyse the mathematical model that deals a standard SIR (Susceptible, Infected, Removed) epidemiological model with prevalence dependent behaviour change as was described in [Kassa and Ouhinou \(2011\)](#).

We begin our exploration by extending the traditional SIR epidemiological model to include a cohort of individuals in a population who take part in consistent self-protective measures because of their awareness about the disease and their perceived risk of getting infected. We shall call such a compartment an ‘*Educated*’ cohort and denote by  $E(t)$ , the number of individuals in this class at any time  $t$ . The individuals in this class have a reduced susceptibility due to their usage of any of the available self-protective measures against the disease. For susceptible individuals to change their risky behaviour and move into the  $E$  class, first, they need to get information about the disease and its protective mechanisms; and they should perceive the risk of contracting the disease. They may acquire this information through their social network (i.e. from people who have infection experience or from among their circle who is informed about it) or through some public health agents or through mass media. Second, they need to be informed about the possible ways of preventing oneself from the disease. Therefore, the reduction in their susceptibility to the infection depends on the efficacy levels of the preventive mechanisms that each individual is using. If we denote the average efficacy of all existing self-protective measures for the disease by a constant  $\gamma$ , the susceptibility level of each individual in  $E$  is reduced by  $100 \times \gamma$  % on average.

To model the flow of individuals from the *Susceptible* ( $S$ ) class to the *Educated* ( $E$ ) class, we propose to view self-protective measures as “*innovation*” and its adoption by the population to follow the so called the “*diffusion*” process, borrowing the term “*Diffusion of Innovations*” from [Rogers \(1983\)](#). It has been accepted by public health researchers since long time (see, for eg., [Green and McAlister 1984](#); [Green et al. 2009](#); [Oldenburg and Glanz 2008](#)) that the adoption of the programmes of public health interventions at the population level follows a logistic curve through time, similar to the diffusion of innovations or ideas in populations. The curve is initiated by the adoption of the so called “*Innovators*”, followed by “*Early adopters*”. The concavity of the curve changes in the middle of its hight when the “*Early majority*” population start to adopt the innovation and it saturates asymptotically to its full hight (100 %) by the time when the “*Late adopters*” participate into the programme ([Green and McAlister 1984](#)).

Moreover, behaviour change is observed to be increasing with increasing prevalence. For instance, in the study by [Philipson \(1996\)](#), demand for measles vaccination increased as prevalence increases in the USA and similarly in [Ahituv et al. \(1996\)](#)

it has been indicated that the use of condoms was quite responsive to the prevalence of AIDS in one's state of residence, and this responsiveness has been increasing over time in the USA. That means, a self preventive behaviour change is prevalence-elastic.

Following the above discussion, it is justifiable to use the same mathematical formulation as in [Kassa and Ouhinou \(2011\)](#) to simulate the flow of individuals from class  $S$  to  $E$ .

We assume that the disease is not curable but has a treatment which reduces the infectiousness level by  $(1 - \epsilon)\%$  and gives a better quality of life to the infected individuals. Thus, in place of the Removed (R) class in the traditional SIR model, we use the *Treated* (T) class to represent the set of infected individuals who are receiving treatments (like the ART in the case of HIV). With these modifications, the dynamics of the epidemic can be formulated by the following system of ordinary differential equations.

$$\begin{aligned} S' &= \pi - \alpha e(P)S - \lambda S - \mu S \\ E' &= \alpha e(P)S - (1 - \gamma)\lambda E - \mu E \\ I' &= \lambda S + (1 - \gamma)\lambda E - \rho I - (\mu + d_1)I \\ T' &= \rho I - (\mu + d_2)T, \end{aligned} \quad (1)$$

where  $N(t) = S(t) + E(t) + I(t) + T(t)$ , and  $\lambda = c\beta \frac{I + \epsilon T}{N}$  represents the force of infection (with  $c$  and  $\beta$  representing the contact and transmission rates, respectively),  $\mu$  represents the natural death rate,  $d_1$  and  $d_2$  represent the rate of disease induced deaths from classes  $I$  and  $T$ , respectively, and  $\rho$  represents the rate at which infected individuals are recruited to receive treatment. Because of the existence of the diseases induced deaths the total population also varies as:

$$N' = \pi - \mu N - d_1 I - d_2 T.$$

Following the previous discussion since  $P = \frac{I + T}{N}$ , we set (as in [Kassa and Ouhinou \(2011\)](#)) the value of the parameter  $e(P)$  (and hence of  $e(t)$ ) to take the form:

$$e(P) = \frac{P^n}{P_*^n + P^n} \quad \text{or equivalently} \quad e(t) = \frac{(I + T)^n}{N^n P_*^n + (I + T)^n}.$$

With regard to the well-posedness and boundedness of solutions of the above system, we refer to [Kassa and Ouhinou \(2011\)](#) which shows that for every initial condition in the positive cone  $\mathcal{O}$  the system (1) has a unique maximal solution on  $[0, t_{max})$ .

### 3 Steady states and their stability

The stability of the equilibrium points of system (1) is governed by the basic reproduction number  $\mathcal{R}_o$ , the expected number of secondary cases produced by a single infected individual in a wholly susceptible population.  $\mathcal{R}_o$  can be determined by using the decomposition technique presented in van den [Driessche and Watmough \(2002\)](#). Thus we find it to be

$$\mathcal{R}_o = \frac{c\beta}{\mu + \rho + d_1} + \frac{\epsilon\rho c\beta}{(\mu + \rho + d_1)(\mu + d_2)}. \tag{2}$$

To calculate the steady state solutions of the system, let the size of the population in each compartment at the steady state be given by the vector  $\mathcal{E} = (S^*, E^*, I^*, T^*)$ . The steady states are obtained by equating each of the time derivatives in (1) to zero. By doing so, we get that

$$T^* = \frac{\rho}{\mu + d_2} I^*.$$

Substituting in the expressions of  $\lambda^*$  and  $P^*$  at the steady state  $\mathcal{E}$ , one has

$$\lambda^* = c\beta \left( 1 + \frac{\epsilon\rho}{\mu + d_2} \right) \frac{I^*}{N^*} \quad \text{and} \quad P^* = \frac{\mu + \rho + d_2}{c\beta(\mu + \epsilon\rho + d_2)} \lambda^*.$$

Let  $K = \frac{\mu + \rho + d_2}{c\beta(\mu + \epsilon\rho + d_2)}$ . Then, the function of the behavior change  $e^*$  at the steady state  $\mathcal{E}$  is given by

$$e^* = \frac{(\lambda^*)^n}{\lambda_0^n + (\lambda^*)^n},$$

where  $\lambda_0 = \frac{P^*}{K}$ . After some calculation, we will come up with the following equilibrium points described in terms of the corresponding force of infection  $\lambda^*$ .

$$\begin{aligned} S^* &= \frac{\pi[\lambda_0^n + (\lambda^*)^n]}{\alpha(\lambda^*)^n + [\lambda^* + \mu][\lambda_0^n + (\lambda^*)^n]} \\ E^* &= \frac{\alpha\pi(\lambda^*)^n}{\{\alpha(\lambda^*)^n + [\mu + \lambda^*][\lambda_0^n + (\lambda^*)^n]\}[\mu + (1 - \gamma)\lambda^*]} \\ I^* &= \frac{\pi - \mu(S^* + E^* + T^*)}{\rho + \mu + d_1} \\ T^* &= \frac{\rho}{\mu + d_2} I^* \\ N^* &= S^* + E^* + I^* + T^* \end{aligned}$$

and  $\lambda^*$  is a non negative real root of the polynomial:

$$Q(\lambda) = \lambda \left[ A_0 + A_1\lambda + A_2\lambda^2 + (A_3 + A_4\lambda + A_5\lambda^2)\lambda^n \right], \tag{3}$$

where,

$$\begin{aligned} A_0 &= (1 - \mathcal{R}_o)\mu^2\lambda_0^n \\ A_1 &= \left[ \frac{\mu}{\mu + \rho + d_1} + (1 - \gamma)(1 - \mathcal{R}_o) \right] \mu\lambda_0^n \end{aligned}$$

$$\begin{aligned}
&= (\mathcal{R}_o - 1)\mu\lambda_o^n \left[ \gamma - \left( 1 - \frac{\mu}{(\mathcal{R}_o - 1)(\mu + \rho + d_1)} \right) \right] \\
A_2 &= \left[ (1 - \gamma) \frac{\mu}{\mu + \rho + d_1} \right] \lambda_o^n \\
A_3 &= \mu\alpha\mathcal{R}_o \left[ \gamma - \left( 1 + \frac{\mu}{\alpha} \right) \left( 1 - \frac{1}{\mathcal{R}_o} \right) \right] \\
A_4 &= \left[ \frac{\mu}{\mu + \rho + d_1} + (1 - \gamma)(1 - \mathcal{R}_o) + \frac{\alpha(1 - \gamma)}{\mu + \rho + d_1} \right] \mu \\
A_5 &= \left[ (1 - \gamma) \frac{\mu}{\mu + \rho + d_1} \right]
\end{aligned} \tag{4}$$

When  $\lambda^* = 0$ , we get the disease free equilibrium  $(S^*, E^*, I^*, T^*) = (\frac{\pi}{\mu}, 0, 0, 0) =: \mathcal{E}_o$ . Where positive roots of  $Q(\lambda)$  will correspond to the endemic equilibrium points.

### 3.1 The disease free equilibrium and its global stability

The disease free equilibrium is given by:

$$\mathcal{E}_o = (S^o, E^o, I^o, T^o) = \left( \frac{\pi}{\mu}, 0, 0, 0 \right).$$

The linear stability of the disease free equilibrium  $\mathcal{E}_o$  is governed by the basic reproduction number  $\mathcal{R}_o$ , which is obtained in Eq. (2). Using this value we get the following results.

**Theorem 1** *The model system (1) always has the disease-free equilibrium  $\mathcal{E}_o$ . If  $\mathcal{R}_o < 1$ , then  $\mathcal{E}_o$  is locally asymptotically stable, and unstable otherwise.*

*Proof* This is an immediate consequence of Theorem 2 in van den Driessche and Watmough (2002).  $\square$

Moreover the global stability is discussed in the following result.

**Theorem 2** *If  $\mathcal{R}_o < 1$ , the disease free equilibrium  $\mathcal{E}_o$  is globally asymptotically stable.*

*Proof* Consider the third equation in the system (1),

$$I'(t) = c\beta \frac{S + (1 - \gamma)E}{N} (I + \epsilon T) - (\mu + \rho + d_1)I(t),$$

for which the solutions are given by the following variation of constants formula

$$I(t) = I(0)e^{-(\mu + \rho + d_1)t} + \int_0^t e^{(\mu + \rho + d_1)(t-s)} c\beta \frac{S + (1 - \gamma)E}{N} (I + \epsilon T) ds.$$

Which implies that

$$\begin{aligned}
 I(t) &\leq I(0)e^{-(\mu+\rho+d_1)t} + \int_0^t e^{(\mu+\rho+d_1)(t-s)} c\beta(I + \epsilon T)ds \\
 &\leq I(0)e^{-(\mu+\rho+d_1)t} + \int_0^{\frac{t}{2}} e^{(\mu+\rho+d_1)(t-s)} c\beta(I + \epsilon T)ds \\
 &\quad + \int_{\frac{t}{2}}^t e^{(\mu+\rho+d_1)(t-s)} c\beta(I + \epsilon T)ds.
 \end{aligned}$$

Since the functions  $I$  and  $T$  are bounded for  $t \geq 0$  by  $\frac{\pi}{\mu}$ , we get

$$\lim_{t \rightarrow \infty} \left[ I(0)e^{-(\mu+\rho+d_1)t} + \int_0^{\frac{t}{2}} e^{(\mu+\rho+d_1)(t-s)} c\beta(I + \epsilon T)ds \right] = 0.$$

Thus,

$$\begin{aligned}
 \limsup_{t \rightarrow \infty} I(t) &\leq \limsup_{t \rightarrow \infty} \left[ c\beta \int_{\frac{t}{2}}^t e^{(\mu+\rho+d_1)(t-s)} ds \sup_{\frac{t}{2} \leq s \leq t} (I(s) + \epsilon T(s)) \right] \\
 &\leq c\beta \limsup_{t \rightarrow \infty} \int_{\frac{t}{2}}^t e^{(\mu+\rho+d_1)(t-s)} ds \limsup_{t \rightarrow \infty} (I + \epsilon T) \\
 &\leq \frac{c\beta}{\mu + \rho + d_1} \limsup_{t \rightarrow \infty} (I + \epsilon T). \tag{5}
 \end{aligned}$$

On the other hand, from the fourth equation of system (1), we have

$$\begin{aligned}
 T(t) &= T(0)e^{(\mu+d_2)t} + \int_0^t e^{(\mu+d_2)(t-s)} \rho I ds \\
 &= T(0)e^{(\mu+d_2)t} + \int_0^{\frac{t}{2}} e^{(\mu+d_2)(t-s)} \rho I ds + \int_{\frac{t}{2}}^t e^{(\mu+d_2)(t-s)} \rho I ds.
 \end{aligned}$$



For which, we have

$$\lim_{t \rightarrow \infty} \left[ T(0)e^{(\mu+d_2)t} + \int_0^{\frac{t}{2}} e^{(\mu+d_2)(t-s)} \rho I ds \right] = 0.$$

Similarly, this implies that

$$\begin{aligned} \limsup_{t \rightarrow \infty} T(t) &\leq \limsup_{t \rightarrow \infty} \int_{\frac{t}{2}}^t e^{(\mu+d_2)(t-s)} \rho I ds \\ &\leq \limsup_{t \rightarrow \infty} \rho \int_{\frac{t}{2}}^t e^{(\mu+d_2)(t-s)} ds \limsup_{t \rightarrow \infty} I(t) \\ &\leq \frac{\rho}{\mu + d_2} \limsup_{t \rightarrow \infty} I(t). \end{aligned} \tag{6}$$

Substituting (5) in (6), we come up with an inequality:

$$\limsup_{t \rightarrow \infty} T(t) \leq \frac{c\beta\rho}{(\mu + \rho + d_1)(\mu + d_2)} \limsup_{t \rightarrow \infty} (I + \epsilon T).$$

Consequently, we get

$$0 \leq \limsup_{t \rightarrow \infty} (I + \epsilon T) \leq \mathcal{R}_0 \limsup_{t \rightarrow \infty} (I + \epsilon T). \tag{7}$$

As a conclusion, if  $\mathcal{R}_0 < 1$ , from (7) it follows that

$$\lim_{t \rightarrow \infty} (I + \epsilon T) = \limsup_{t \rightarrow \infty} (I + \epsilon T) = 0.$$

This completes the proof. □

Since the effect of behaviour change could be observed when the number of infected individuals get to the level that the population can observe it, the disease free equilibrium will not be affected by the effects of behaviour change. This will in turn forces the value of  $\mathcal{R}_o$  to remain the same as in the case when the effect of learning is not taken into consideration in the model. If  $\mathcal{R}_o$  is small (i.e, if it is  $< 1$ ) the disease will die out and the prevalence does not come closer to  $P_*$ . In such cases, the reaction of the population against the disease will be minimum. However, if  $\mathcal{R}_o > 1$  the prevalence increases as time goes and due to this fact we may see a significant effect in the change of behaviour in protecting oneself from the disease. This suggests that the effect of learning or behaviour modification as self protective measures will have more significance on the dynamics of the disease in the endemic case.

### 3.2 Endemic equilibrium and persistence of the disease

Each of the non zero positive real roots of the polynomial (3) will determine the endemic equilibrium points of the system (1). Since  $Q(\lambda)$  is an  $(n + 2)^{\text{th}}$  degree polynomial it is difficult to explicitly calculate its roots. However, since we are interested only on the positive real roots of the polynomial we will qualitatively study the behaviour of the roots from the sign of the coefficients  $A_0, \dots, A_5$  as it was done in Kassa and Ouhinou (2011). The signs of these coefficients will be controlled by the choice of the parameters  $\alpha$ , the rate of dissemination of convincing information into the population, and  $\gamma$ , the average efficacy level of the various self protective actions to be taken by the population.

From the expressions in (4) it is clear that for all values of  $\alpha$  and  $\gamma$  in  $[0, 1]$ , if  $\mathcal{R}_o < 1$  then all the coefficients of  $Q(\lambda)$  are positive. But if  $\mathcal{R}_o > 1$ ,  $A_0 < 0$  and  $A_2, A_5 > 0$ . Therefore, the number of sign changes in the coefficients  $A_0, A_1, A_2, A_3, A_4$  and  $A_5$  are exactly one or three. Using Descartes' rule of signs on the polynomial  $Q(\lambda)$ , we conclude that  $Q(\lambda)$  has either one or three positive roots. For specific details we need to determine the signs of the coefficients  $A_1, A_3$  and  $A_4$ . One can see that  $A_1 > 0$  implies  $A_4 > 0$ .

We will now see different cases where the signs of the three coefficients vary. If we set

$$\gamma_0 := 1 - \frac{\mu}{(\mathcal{R}_o - 1)(\mu + \rho + d_1)}$$

One can easily see that  $A_1$  is positive for all  $\gamma \in [\gamma_0, 1]$ .

On the other hand the graph of the function:

$$\gamma_4(\alpha) = 1 + \frac{\mu}{(1 - \mathcal{R}_o)(\mu + \rho + d_1) + \alpha}$$

separates between the positive and negative values of  $A_4$ .

Moreover, the sign of coefficient  $A_3$  is determined by the graph of

$$\gamma_3(\alpha) = \left(1 - \frac{1}{\mathcal{R}_o}\right) \left(1 + \frac{\mu}{\alpha}\right).$$

Thus we have the following theorem:

**Theorem 3** For  $\mathcal{R}_o > 1$ , the system (1) has exactly one or three endemic state equilibrium points. Moreover, the system has a unique endemic state equilibrium if either one of the following conditions hold.

- (i)  $1 < \mathcal{R}_o \leq 1 + \frac{\mu}{\mu + \rho + d_1}$  and  $\mu(\mathcal{R}_o - 1) \leq \alpha \leq 1$  with  $(1 - \frac{1}{\mathcal{R}_o})(1 + \frac{\mu}{\alpha}) \leq \gamma \leq 1$ .
- (ii)  $1 + \frac{\mu}{\mu + \rho + d_1} < \mathcal{R}_o \leq 1 + \frac{1}{\mu}$  and  $\mu(\mathcal{R}_o - 1) \leq \alpha \leq 1$  with  $\max \left\{ 1 + \frac{\mu}{(\mu + \rho + d_1)(1 - \mathcal{R}_o)}, (1 - \frac{1}{\mathcal{R}_o})(1 + \frac{\mu}{\alpha}) \right\} \leq \gamma \leq 1$ .
- (iii)  $1 + \frac{\mu}{\mu + \rho + d_1} < \mathcal{R}_o \leq 1 + \frac{1}{\mu}$  and  $\frac{\mu(\mu + \rho + d_1)(\mathcal{R}_o - 1)^2}{d_1(\mathcal{R}_o - 1) - \mu} \leq \alpha \leq 1$  with  $(1 - \frac{1}{\mathcal{R}_o})(1 + \frac{\mu}{\alpha}) \leq \gamma \leq 1 + \frac{\mu}{(\mu + \rho + d_1)(1 - \mathcal{R}_o)}$ .

In the above theorem, we could only prove the uniqueness of the endemic state equilibrium (for  $\mathcal{R}_o > 1$ ) when the values of  $\alpha$  and  $\gamma$  satisfy the given conditions. However, we couldn't prove or disprove the uniqueness of the equilibrium in the remaining cases. As was expressed in [Kassa and Ouhinou \(2011\)](#), however, numerical results suggest the existence of only one endemic state equilibrium point paralleling what is reported in [Chen \(2004\)](#).

To establish the local stability of system (1) around  $\mathcal{R}_o = 1$ , we use the center manifold theorem from [Castillo-Chavez and Song \(2004\)](#). For this purpose we introduce  $x_1 = \frac{\pi}{\mu} - S$ ,  $x_2 = E$ ,  $x_3 = I$  and  $x_4 = T$ . As in other studies, we can see that  $\mathcal{R}_0$  is linearly related to the rate of infection  $c\beta$ . So, we choose  $\phi = c\beta$  to be the bifurcation parameter around  $\mathcal{R}_o = 1$ , which corresponds to  $\phi = \phi^* = \frac{(\mu + \rho + d_1)(\mu + d_2)}{\epsilon\rho + \mu + d_2}$ . Using these notations, system (1) will take the form:

$$\begin{aligned}
 f_1 &:= x'_1(t) = \alpha e(P)\left(\frac{\pi}{\mu} - x_1\right) + \phi \frac{\left(\frac{\pi}{\mu} - x_1\right)(x_3 + \epsilon x_4)}{N} - \mu x_1 \\
 f_2 &:= x'_2(t) = \alpha e(P)\left(\frac{\pi}{\mu} - x_1\right) - (1 - \gamma)\phi \frac{x_2(x_3 + \epsilon x_4)}{N} - \mu x_2 \\
 f_3 &:= x'_3(t) = \phi \frac{\left(\frac{\pi}{\mu} - x_1 + (1 - \gamma)x_2\right)(x_3 + \epsilon x_4)}{N} - (\mu + \rho + d_1)x_3, \\
 f_4 &:= x'_4(t) = \rho x_3 - (\mu + d_2)x_4
 \end{aligned}
 \tag{8}$$

where  $N = \frac{\pi}{\mu} - x_1 + x_2 + x_3 + x_4$ . It is clear that the disease free equilibrium corresponds to  $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*) = (0, 0, 0, 0)$ . Now that the linearisation matrix of (8) at  $\mathbf{x}^*$  is given by

$$D_{\mathbf{x}}\mathbf{f} = \begin{pmatrix} -\mu & 0 & \phi^* & \epsilon\phi^* \\ 0 & -\mu & 0 & 0 \\ 0 & 0 & \phi^* - (\mu + \rho + d_1) & \epsilon\phi^* \\ 0 & 0 & \rho & -(\mu + d_2) \end{pmatrix}$$

So, when  $\phi < \phi^*$  all the eigenvalues of this matrix have negative real parts. For  $\phi = \phi^*$  the matrix has a zero as its simple eigenvalue, and the remaining eigenvalues are negative real numbers. And when  $\phi > \phi^*$ , the matrix has one positive eigenvalue which is simple and all the remaining ones are negative. Thus for  $\mathcal{R}_o > 1$ , the disease free equilibrium  $\mathcal{E}_o$  is a saddle point of the system (1) with  $\dim W^s(\mathcal{E}_o) = 2$  and  $\dim W^u(\mathcal{E}_o) = 1$ , where  $W^s(\mathcal{E}_o)$  and  $W^u(\mathcal{E}_o)$  are, respectively, the stable and unstable manifolds of the system (1).

To determine the kind of bifurcation we have at  $\mathcal{R}_o = 1$ , when  $\phi$  crosses  $\phi^*$  from left to right, we apply Theorem 4.1 in [Castillo-Chavez and Song \(2004\)](#). To this end, for  $\phi = \phi^*$ , the corresponding right- and left-eigenvectors of  $D_{\mathbf{x}}\mathbf{f}$  associated with eigenvalue 0 are respectively  $\mathbf{w} = \left(\frac{\phi^*}{\mu} \left(1 + \frac{\epsilon\rho}{\mu+d_2}\right), 0, 1, \frac{\rho}{\mu+d_2}\right)^T$  and  $\mathbf{v} = \left(0, 0, 1, \frac{\epsilon\phi^*}{\mu+d_2}\right)$ . The second derivatives of  $\mathbf{f} = (f_1, f_2, f_3, f_4)$  corresponding to the non-zero components of  $\mathbf{w}$  and  $\mathbf{v}$  evaluated at  $\mathbf{x}^*$  are

$$\begin{aligned} \frac{\partial^2 f_3}{\partial x_1^2} &= 0, & \frac{\partial^2 f_3}{\partial x_3 \partial x_1} &= 0, & \frac{\partial^2 f_3}{\partial x_4 \partial x_1} &= 0, \\ \frac{\partial^2 f_3}{\partial x_1 \partial x_3} &= 0, & \frac{\partial^2 f_3}{\partial x_3^2} &= -2\phi^* \frac{\mu}{\pi}, & \frac{\partial^2 f_3}{\partial x_4 \partial x_3} &= -(1 - \epsilon)\phi^* \frac{\mu}{\pi}, \\ \frac{\partial^2 f_3}{\partial x_1 \partial x_4} &= 0, & \frac{\partial^2 f_3}{\partial x_3 \partial x_4} &= -(1 + \epsilon)\phi^* \frac{\mu}{\pi}, & \frac{\partial^2 f_3}{\partial x_4^2} &= -2\epsilon\phi^* \frac{\mu}{\pi}, \\ \frac{\partial^2 f_3}{\partial x_1 \partial \phi} &= 0, & \frac{\partial^2 f_3}{\partial x_3 \partial \phi} &= 1, & \frac{\partial^2 f_3}{\partial x_4 \partial \phi} &= \epsilon \end{aligned}$$

and

$$\frac{\partial^2 f_4}{\partial x_i \partial x_j} = 0, \quad \frac{\partial^2 f_4}{\partial x_i \partial \phi} = 0 \quad \text{for } i, j \in \{1, 3, 4\}$$

We now need to compute the following values:

$$\begin{aligned} a &= \sum_{k,j,i=1}^3 v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(\mathbf{x}^*, \phi^*) & b &= \sum_{k,i=1}^3 v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi}(\mathbf{x}^*, \phi^*) \\ &= -\frac{\phi^* \pi}{\mu} \left(1 + \frac{\epsilon \phi^*}{\mu + d_2}\right) \left(1 + \frac{\rho}{\mu + d_2} + \epsilon \frac{\rho^2}{(\mu + d_2)^2}\right) & &= \left(1 + \frac{\epsilon \phi^*}{\mu + d_2}\right) \left(1 + \frac{\epsilon \rho}{\mu + d_2}\right) \\ &< 0 & &> 0. \end{aligned}$$

Since  $a < 0$  and  $b > 0$ , thus we have verified the statement that when  $\phi$  passes through the point  $\phi^*$  from left to right, the stability of  $\mathcal{E}_o = (\frac{\pi}{\mu}, 0, 0, 0)$  changes from stable to unstable, correspondingly, a negative unstable equilibrium becomes positive and locally asymptotically stable. The following theorem summarizes the above results.

**Theorem 4** *If  $\mathcal{R}_o < 1$ , the system (1) has a unique biologically feasible equilibrium point  $\mathcal{E}_o$  which is a disease-free equilibrium and is globally asymptotically stable.*

*If  $\mathcal{R}_o > 1$ , the disease-free equilibrium  $\mathcal{E}_o$  becomes a saddle point with  $\dim W^s(\mathcal{E}_o) = 2$  and  $\dim W^u(\mathcal{E}_o) = 1$ , where  $W^s(\mathcal{E}_o)$  and  $W^u(\mathcal{E}_o)$  are, respectively, the stable and unstable manifolds of the system (1).*

*When  $\phi$  crosses  $\phi^*$  from left to right (near  $\phi^*$ ), the disease-free equilibrium changes its stability from globally asymptotically stable to unstable. Correspondingly, a unique positive endemic equilibrium  $E^*$  appears which is locally asymptotically stable.*

*Moreover, for  $\phi$  large enough, the system has either one or three positive endemic equilibria.*

### 4 The optimal control model

#### 4.1 Formulation of the control

The possible interventions for any kind of disease, that has a treatment which may not completely cure the disease, can be categorized in to three classes: preventive education, refining and validating the effectiveness of protective mechanisms, and treating the infected individuals. In the sequel these interventions will serve as control parameters in the dynamics of the epidemic model.

- (a) *Educating the population* As it is mentioned in the introduction part, the population can naturally react against the spread of a disease and moreover, if the disease has been known to the population for some time, we assume that the current level of preventive education campaigns by various agents have convinced upto  $100 \times (\alpha_o \times e)\%$  (for some  $\alpha_o > 0$ ) of the susceptible population per year to effectively participate in the self protective schemes available to them. But since the first few proportion of the population can be convinced easily to use the self-protective measures while any additional proportion requires more effort and higher cost, we may need a cost adjustment value in the objective function to take this into account. Thus, the control in this category results in increasing (or decreasing) from its current value in the rate of information dissemination that can possibly convince susceptible individuals to participate in modifying their risky behaviour. Assume that the control function  $u_1(t)$  measures the rate at which additional susceptible individuals are convinced to take part in behaviour modification. Its application in the dynamics is modelled by simply replacing the term  $\alpha$  in (1) by  $(\alpha_o + u_1(t))$ . However, the cost of the effort in convincing the population for behaviour modification becomes expensive as the proportion of the non-convinced susceptible individuals gets smaller (Green and McAlister 1984).

This situation can be captured by including the term  $\left(\frac{E}{N}\right)^m$ , where  $m$  is a constant positive integer, as part of the coefficient for  $u_1^2(t)$  in the objective function (see, for example, Behncke 2000; Gaff and Schaefer 2009 for inclusion of such factors). Here we assume that the larger the proportion of the ‘*Educated*’ class, the lower will be the proportion of individuals in the susceptible population. In Gaff and Schaefer (2009) it is indicated that taking  $m = 10$  has resulted in more realistic numerical simulation values for a model with vaccination intervention. Because of practicality and economical limitations on the maximum rate of convincing individuals for behaviour modification, we also assume that  $\alpha_{\max} > 0$  to be the maximum rate.

- (b) *Improve the average efficacy of the protective measures* This intervention could be realized in terms of investing on activities that could find alternative measures in self protective mechanisms available to be used by the ‘*Educated*’ class with better efficacy levels, improve the existing measures so that the efficacious once could be chosen and constantly used by more individuals, and/or by making the means (or devices) of self protective mechanisms available to every one in affordable price. Intuitively it is clear to conclude that the more efficacious the protective mechanisms are the less number of new infections will result from among the ‘*Educated*’ class, and hence the more individuals could be persuaded to decide in using the protective mechanisms and adhere to them. Therefore, if we assume that  $u_2(t)$  represents the total amount of additional efforts (in percentage) made to increase the efficacy of the protective mechanisms from its current level  $\gamma_o$ , then the total number of new infections from the ‘*Educated*’ class per unit time, after this effort can be formulated in the dynamics as

$$(1 - \gamma_o - u_2)\lambda E.$$

That means, when there is no additional intervention we still have the current rate of new infection. But when the intervention reaches its maximum, 1 (i.e., when  $\gamma_o + u_2(t) = 1$ ), we will achieve a 100 % protection, i.e., there will be no new infection arising from the educated class, which is the ideal case.

- (c) *Validate the rate of recruitment for treatment* Treatment has epidemiological advantage to the general population that the rate of disease transmission will be suppressed by a certain level (for instance, a 92 % reduction in transmission rate of HIV is reported in [Donnell et al. \(2010\)](#)), and to the infected individual by increasing the level of his/her quality of life. We assume that the cost of treatment is shared by the entire population, and the analysis will be made at population level rather than individual level. If we assume that the control function  $u_3(t)$  measures the rate at which additional infectious individuals transform to the ‘Treated’ class at any time  $t$ , where the current rate is at  $\rho_0$ , this control will be seen in the dynamics as  $(\rho_0 + u_3(t))I(t)$  by replacing  $\rho I(t)$  in (1). Again as in case (a) above, due to economical and logistic reasons, there are limitations on the maximum rate at which individuals are recruited to get treatment at each time period. Thus, let the constant  $\rho_{\max} > 0$  represent this maximum rate.

Using the above described control parameters, the system of the disease dynamics can be rewritten as:

$$\begin{aligned}
 \frac{dS}{dt} &= \pi - (\alpha_o + u_1)eS - \lambda S - \mu S \\
 \frac{dE}{dt} &= (\alpha_o + u_1)eS - (1 - \gamma_o - u_2)\lambda E - \mu E \\
 \frac{dI}{dt} &= \lambda S + (1 - \gamma_o - u_2)\lambda E - (\rho_0 + u_3)I - (\mu + d_1)I, \\
 \frac{dT}{dt} &= (\rho_0 + u_3)I - (d_2 + \mu)T,
 \end{aligned}
 \tag{9}$$

where  $\lambda = c\beta \frac{I + \epsilon T}{N}$ ,  $u_1(t) \in [-\alpha_o, \alpha_{\max} - \alpha_o]$ ,  $u_2(t) \in [-\gamma_o, 1 - \gamma_o]$ ,  $u_3(t) \in [-\rho_o, \rho_{\max} - \rho_o]$  for all  $t \in [0, t_f]$ .

Thus, given initial population size of each cohort  $S_0, E_0, I_0$  and  $T_0$ , our main goal here is to minimise the total number of new infections in the planning period, while also minimizing the total cost of controlling the disease dynamics. That means, by constructing optimal values of Lebesgue integrable, bounded control functions  $u_i(t)$ ,  $i = 1, 2, 3$ , we seek the best strategy that can control the dynamics of the epidemics modelled in Eq. (9). To this end, we minimize the objective functional

$$J(u_1, u_2, u_3) = \int_0^{t_f} \left( A_1 I(t) + A_2 T(t) + \frac{B_1}{2} \left( \frac{E}{N} \right)^m u_1^2(t) + \frac{B_2}{2} u_2^2(t) + \frac{B_3}{2} u_3^2(t) \right) dt,
 \tag{10}$$

where  $u_1, u_2$  and  $u_3$  are Lebesgue measurable bounded functions on  $[0, t_f]$ .

Since implementation of any public health intervention has increasing costs with reaching higher fraction of the population, we usually take a non-linear cost function, like the quadratic. The constants  $A_1, A_2, B_1, B_2$  and  $B_3$  could be considered as values that will balance the units of measurement and also may indicate the importance of one type of intervention over the other, at implementation level to the general public.

#### 4.2 Existence and characterization of optimal control solution

The first task will be to examine conditions that can assure the existence of a solution to our optimal control problem.

**Theorem 5** (Existence of optimal control solution) *There exists an optimal control triple  $\mathbf{u}^* = (u_1^*, u_2^*, u_3^*)$ , and corresponding solution vector  $\mathbf{x}^* = (S^*, E^*, I^*, T^*)$  to the state initial value problem (9) that maximizes the objective functional  $J(\mathbf{u})$  of (10) over the set of admissible controls  $\mathcal{U}$ .*

*Proof* Let  $f(\mathbf{x}, \mathbf{u}, t) = A_1 I(t) + A_2 T(t) + \frac{B_1}{2} \left(\frac{E}{N}\right)^m u_1^2(t) + \frac{B_2}{2} u_2^2(t) + \frac{B_3}{2} u_3^2(t)$ ,

$$g(\mathbf{x}, \mathbf{u}, t) = \frac{d\mathbf{x}}{dt} = \begin{cases} \frac{dS}{dt} = \pi - (\alpha_o + u_1)eS - c\beta \frac{I + \epsilon T}{N} S - \mu S \\ \frac{dE}{dt} = (\alpha_o + u_1)eS - (1 - \gamma_o - u_2)c\beta \frac{I + \epsilon T}{N} E - \mu E \\ \frac{dI}{dt} = c\beta \frac{I + \epsilon T}{N} S + (1 - \gamma_o - u_2)\lambda E - (\rho_0 + u_3)I - (\mu + d_1)I \\ \frac{dT}{dt} = (\rho_0 + u_3)I - (d_2 + \mu)T, \end{cases}$$

and  $\mathcal{U} = \{(u_1(t), u_2(t), u_3(t)) \in L^1(0, t_f) \mid -\alpha_o \leq u_1(t) \leq \alpha_{\max} - \alpha_o, -\gamma_o \leq u_2(t) \leq 1 - \gamma_o, -\rho_o \leq u_3(t) \leq \rho_{\max} - \rho_o, \forall t \in [0, t_f]\}$ . □

Since all the involved functions in our model are continuously differentiable, we need to verify the following four conditions given in Filippov-Cesari Theorem (cf. Theorem 3.1 in Hartl et al. 1995).

1. There exists an admissible solution pair  $(\mathbf{x}, \mathbf{u})$ .
2. Roxin’s condition holds, i.e,

$$\Gamma(\mathbf{x}, t) = \{(f(\mathbf{x}, \mathbf{u}, t) + \xi, g(\mathbf{x}, \mathbf{u}, t)) \mid \xi \leq 0, \mathbf{u} \in \mathcal{U}\} \subseteq \mathbb{R}^5$$

is convex for all  $(\mathbf{x}, t) \in \mathbb{R}^4 \times [0, T]$ .

3. There exist  $\delta > 0$  such that  $\|x\| < \delta$  for all admissible  $\{\mathbf{x}(t), \mathbf{u}(t)\}$  and  $t$ .
4. There exist  $\delta_1 > 0$  such that  $\|\mathbf{u}\| < \delta_1$  for all  $\mathbf{u} \in \mathcal{U}(\mathbf{x}, t)$  with  $\|\mathbf{x}\| < \delta$ .

With regard to the first condition, the bound established for the non-controlled dynamics (1) has the same form if we also incorporate the control function parameters, as they eventually add up to zero. Hence, for any  $\mathbf{u} \in \mathcal{U}$  and the state variables, we have

$$0 \leq N(t) \leq \frac{\pi}{\mu}. \tag{11}$$

Moreover, the state system is continuous and bounded for any admissible control  $\mathbf{u} \in \mathcal{U}$ . Therefore, the state system (9) has a unique solution corresponding to every

admissible control  $\mathbf{u} \in \mathcal{U}$ . [see Theorem I.3.1 in [Coddington and Levinson \(1972\)](#), Theorem 9.2.1 in [Lukes \(1982\)](#).]

The state system (9) is linear with respect to the control variables and  $\mathcal{U}$  is compact, which clearly implies Roxin’s condition. Conditions 3) and 4) follow from (11) and the definition of the control set  $\mathcal{U}$ .

Therefore, by Filippov-Cesari Theorem, there exists an optimal control pair  $\{\mathbf{x}^*, \mathbf{u}^*\}$  with  $\mathbf{u}^*(\cdot)$  measurable, that solves the optimal control problem (10).

To formulate necessary conditions for optimality, we need to define the Hamiltonian function of the optimal control problem, which is given by

$$\begin{aligned}
 H(\mathbf{x}, \mathbf{u}, \lambda, t) = & \left( A_1 I + A_2 T + \frac{B_1}{2} \left( \frac{E}{N} \right)^m u_1^2 + \frac{B_2}{2} u_2^2 + \frac{B_3}{2} u_3^2 \right) \\
 & + \lambda_1 \left( \pi - (\alpha_o + u_1) eS - c\beta \frac{I + \epsilon T}{N} S - \mu S \right) \\
 & + \lambda_2 \left( (\alpha_o + u_1) eS - (1 - \gamma_o - u_2) c\beta \frac{I + \epsilon T}{N} E - \mu E \right) \\
 & + \lambda_3 \left( c\beta \frac{I + \epsilon T}{N} S + (1 - \gamma_o - u_2) c\beta \frac{I + \epsilon T}{N} E \right. \\
 & \quad \left. - (\rho_0 + u_3) I - (\mu + d_1) I \right) \\
 & + \lambda_4 ((\rho_0 + u_3) I - (\mu + d_2) T), \tag{12}
 \end{aligned}$$

where,  $\mathbf{x} = (S, E, I, T)$ ,  $\mathbf{u} = (u_1, u_2, u_3)$ ,  $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$

If  $(u_1^*, u_2^*, u_3^*)$  is the optimal control triple yet to be determined, from Pontryagin’s minimum principle (Pontryagin et al. 1962) we have:

1. the minimum conditions, when it occurs in the interior of the control regions:

$$\begin{aligned}
 \frac{\partial H}{\partial u_i} = 0, i = 1, 2, 3. \Rightarrow & \begin{cases} B_1 \left( \frac{E}{N} \right)^m u_1 - \lambda_1 eS + \lambda_2 eS & = 0 \\ B_2 u_2 + \lambda_2 c\beta E \frac{I + \epsilon T}{N} - \lambda_3 c\beta E \frac{I + \epsilon T}{N} & = 0 \\ B_3 u_3 - \lambda_3 I + \lambda_4 I & = 0 \end{cases} \\
 \Rightarrow & \begin{cases} u_1(t) = \frac{1}{B_1} (\lambda_1 - \lambda_2) \left( \frac{E}{N} \right)^{-m} eS \\ u_2(t) = \frac{1}{B_2} (\lambda_3 - \lambda_2) c\beta E \frac{I + \epsilon T}{N} \\ u_3(t) = \frac{1}{B_3} (\lambda_3 - \lambda_4) I \end{cases} \tag{13}
 \end{aligned}$$

2. The adjoint equations:

$$\begin{aligned}
 \dot{\lambda}_1(t) &= - \frac{\partial H}{\partial S} \\
 &= \frac{B_1}{2} \frac{m}{N} \left( \frac{E}{N} \right)^m u_1^2
 \end{aligned}$$



$$\begin{aligned}
 & +(\alpha_o + u_1) \left[ \frac{(I + T)^n \left( N^{n-1} P_*^n (N - nS) + (I + T)^n \right)}{(N^n P_*^n + (I + T)^n)^2} \right] (\lambda_1 - \lambda_2) \\
 & + c\beta(I + \epsilon T) \frac{E + I + T}{N^2} (\lambda_1 - \lambda_3) + (1 - \gamma_o - u_2) c\beta \frac{I + \epsilon T}{N^2} (\lambda_3 - \lambda_2) + \mu\lambda_1;
 \end{aligned} \tag{14}$$

$$\begin{aligned}
 \dot{\lambda}_2(t) &= -\frac{\partial H}{\partial E} \\
 &= -\frac{B_1}{2} \frac{m(N - E)}{N^2} \left( \frac{E}{N} \right)^{m-1} u_1^2 \\
 &+ (\alpha_o + u_1) S \left[ \frac{nN^{n-1} P_*^n (I + T)^n}{(N^n P_*^n + (I + T)^n)^2} \right] (\lambda_2 - \lambda_1) \\
 &+ c\beta S \frac{I + \epsilon T}{N^2} (\lambda_3 - \lambda_1) + (1 - \gamma_o - u_2) c\beta \frac{(I + \epsilon T)(S + I + T)}{N^2} (\lambda_2 - \lambda_3) + \mu\lambda_2;
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 \dot{\lambda}_3(t) &= -\frac{\partial H}{\partial I} \\
 &= \left( \frac{B_1}{2} \frac{m}{N} \left( \frac{E}{N} \right)^m u_1^2 - A_1 \right) \\
 &+ (\alpha_o + u_1) S \left[ \frac{nN^{n-1} P_*^n (I + T)^n (S + E)}{(N^n P_*^n + (I + T)^n)^2} \right] (\lambda_1 - \lambda_2) \\
 &+ c\beta S \frac{S + E + (1 - \epsilon)T}{N^2} (\lambda_1 - \lambda_2) + (1 - \gamma_o - u_2) c\beta E \frac{S + E + (1 - \epsilon)T}{N^2} (\lambda_2 + \lambda_3) \\
 &+ (\rho_o + u_3) (\lambda_3 - \lambda_4) + (\mu + d_1) \lambda_3;
 \end{aligned} \tag{16}$$

$$\begin{aligned}
 \dot{\lambda}_4(t) &= -\frac{\partial H}{\partial T} \\
 &= \left( \frac{B_1}{2} \frac{m}{N} \left( \frac{E}{N} \right)^m u_1^2 - A_2 \right) + (\alpha_o + u_1) S \left[ \frac{nN^{n-1} P_*^n (I + T)^n (S + E)}{(N^n P_*^n + (I + T)^n)^2} \right] (\lambda_1 - \lambda_2) \\
 &+ c\beta S \frac{\epsilon(S + E) + (\epsilon - 1)I}{N^2} (\lambda_1 - \lambda_3) \\
 &+ (1 - \gamma_o - u_2) c\beta E \frac{\epsilon(S + E) + (\epsilon - 1)I}{N^2} (\lambda_2 + \lambda_3) + (\mu + d_2) \lambda_4.
 \end{aligned} \tag{17}$$

3. The transversality conditions:

$$\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = 0 \tag{18}$$

Moreover, from the conditions that  $u_1(t) \in [-\alpha_o, \alpha_{\max} - \alpha_o]$ ,  $u_2(t) \in [-\gamma_o, 1 - \gamma_o]$ ,  $u_3(t) \in [-\rho_o, \rho_{\max} - \rho_o]$ , for all  $t \in [0, t_f]$  we arrive at:

$$\begin{aligned}
 u_1^* &= \min \left\{ \alpha_{\max}, \max \left\{ 0, \frac{1}{B_1} (\lambda_1 - \lambda_2) \left( \frac{E}{N} \right)^{-m} eS \right\} \right\} \\
 u_2^* &= \min \left\{ 1, \max \left\{ 0, \frac{1}{B_2} (\lambda_3 - \lambda_2) c\beta E \frac{I + \epsilon T}{N} \right\} \right\} \\
 u_3^* &= \min \left\{ \rho_{\max}, \max \left\{ 0, \frac{1}{B_3} (\lambda_3 - \lambda_4) I \right\} \right\}
 \end{aligned} \tag{19}$$

Since the model functions are convex with respect to the control variables, and due to *a priori* boundedness of the state and adjoint (or co-state) functions the optimal solution so obtained is unique for small time  $t_f$  (cf. Fister et al. 1998; Gaff and Schaefer 2009).

#### 4.3 Numerical results and Simulations

We carry out simulations to determine the optimal proportions of combinations of the three different interventions that are used to control the epidemic and the sensitivity of the optimality system with initial values of the parameters. The optimality system described in equations (9) and (14)–(18) is solved using the iterative scheme with Runge-Kutta method of order 4. The solution method mainly follows the algorithm derived by Hackbush (1978) and is described as follows:

1. Guess the control variables, in our case we took  $u_i = 0, i = 1, 2, 3$ .
2. Solve for the state variables from system (9) by applying the new control variables starting from the initial conditions; forward in time.
3. Find the values of  $\lambda(t_f)$  from the transversality conditions (18).
4. Using the state variables calculated in step 2, solve the adjoint system (14)–(17) backward in time from  $\lambda(t_f)$ .
5. Calculate the new control values from the minimum condition (13).
6. Check the stopping criteria for convergence and stop if satisfied. If not satisfied, go to step 2.

In order to validate the parameters and formulate our discussion in a more realistic way, we considered the epidemiological information of HIV, as our model also best suits such kind of disease dynamics. The parameters used for the numerical run are for HIV disease dynamics in Botswana.

For the simulations presented in this section we used the parameter values listed in Table 1. The values for the weight parameters for the objective functional are estimated as follows; we took  $A_1 = 1$  per human and the weight of treated people is taken to be  $A_2 = 1.64$ .<sup>1</sup> The weights of the control functions are taken to be  $B_1 = 6 * (2 * 100)$ , for the weight of the effort in disseminating convincing information for the population to change their behaviour, while  $B_2 = 2 * (2 * 100)$  for the effort of improving the average efficacy of the self protective measures to be used, and  $B_3 = 4 * (2 * 100)$ , for the weight of the effort to increase the rate of recruitment of infected population for treatment, respectively. Here we assumed that treatment is more costly as compared to the implementation of self protective measures put all together. The multipliers (100) for each of the  $B_i$ 's are values to balance the units used. But no significant difference is observed in the optimality system if cost values are slightly modified.

The initial conditions for the state variables are estimated as follows. We assumed that the total adult population in Botswana is 253, 724 with a total of 192, 679 susceptible. Out of the susceptible population we assumed that 20 % are convinced to use any one of the existing self protective measures in that year. Thus, we have  $S_0 = 154, 143$

<sup>1</sup> The ministry of Health of Botswana has calculated the cost of medications, testing and counselling at about USD 600 per person per year. Thus we took  $600/365 \approx 1.64$  per person per day.

**Table 1** Descriptions and values of parameters used in the model

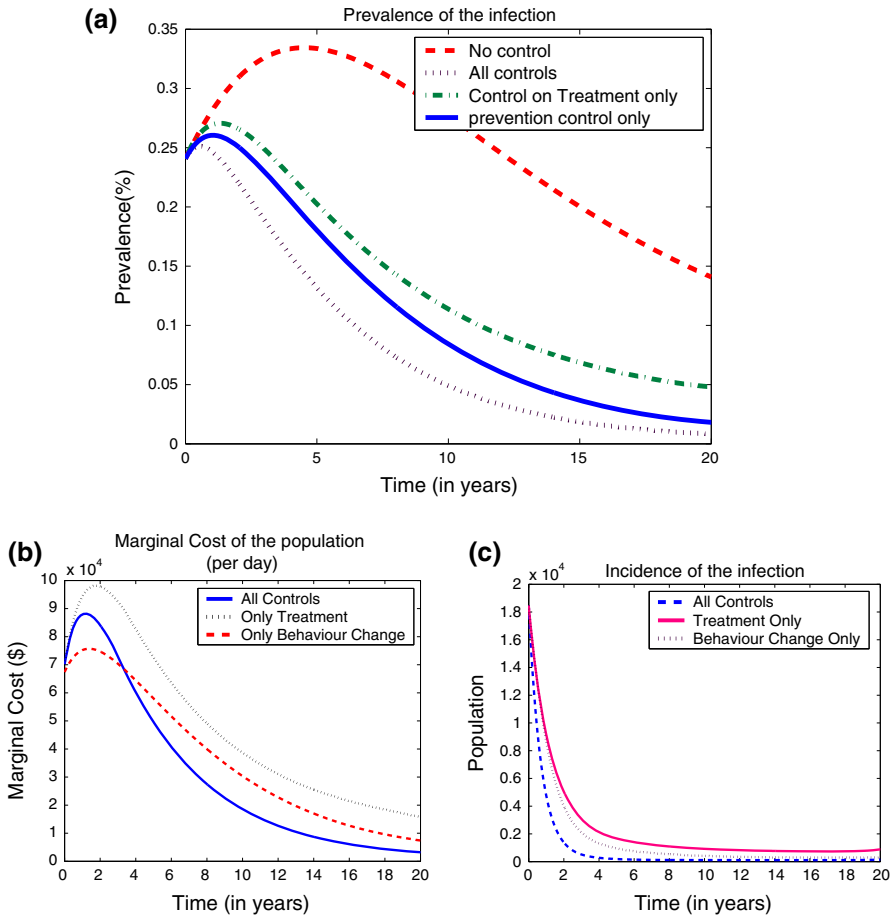
Parameter	Description	Value	Source
$\pi$	Rate of recruitment for susceptible class per year	10,325.25	Kassa and Ouhinou (2011)
$c$	Rate of contact per person	4.87	Kassa and Ouhinou (2011)
$\beta$	Probability of disease transmission	0.12	Vardavas and Blower (2007)
$\epsilon$	Rate of infectiousness of an individual who is receiving a treatment	0.08	Donnell et al. (2010)
$\gamma_o$	Current average efficacy level of the existing protective measures	0.3	(estimate)
$\rho_o$	Current rate of recruitment for treatment per year	0.16 year <sup>-1</sup>	World Health Organization (2005)
$\mu$	Natural death rate for all classes	1/34 year <sup>-1</sup>	Vardavas and Blower (2007)
$d_1$	Disease induced death rate for individuals without treatment	0.1 year <sup>-1</sup>	Vardavas and Blower (2007)
$d_2$	Disease induced death rate for individuals on treatment	0.18 year <sup>-1</sup>	Vardavas and Blower (2007)
$\alpha_o$	Current rate of convincing information dissemination to take part in self-protective actions	0.18 year <sup>-1</sup>	Kassa and Ouhinou (2011)
$n$	Level of reaction of the population for the disease	2	Kassa and Ouhinou (2011)

and  $E_0 = 38,536$ . Out of the total infected population, 61,045, we assumed that only 16 % is put under treatment in the initial year as per the [World Health Organization \(2005\)](#) estimate. This yields, the estimate of  $I_0 = 51,278$  and  $T_0 = 9,767$ .

Moreover, the maximum proportion of population to be convinced to take part in permanent self protective actions against the disease is taken to be  $\alpha_{\max} = 80\%$  (of the susceptible population) per year and the maximum rate of recruitment for treatment is assumed to be  $\rho_{\max} = 90\%$  (of the infected population) per year, while the maximum average efficacy to be reached is taken to be 95 %.

In our numerical runs, we compared optimal combinations of various interventions by varying the possible intervention strategies and taking each control at a time by using the parameter values listed in [Table 1](#) and other values listed above.

In the simulation, we first used the controls  $u_1$  and  $u_2$  (values controlling preventive mechanisms) to optimize the objective function  $\mathbf{J}$ , while the control  $u_3$  on treatment is set to zero. Next we set the controls  $u_1$  and  $u_2$  to zero and optimize the objective functional over the control  $u_3$ . Finally, we optimize the objective functional over all the three control variables. In [Fig. 1a](#), we can observe that the prevalence of the disease increases and stays very high if there is no additional control measure employed. On the other hand, if the strategy focusses only on behavioural change measures, with no additional effort is made to treat more infected people, the result seems better than the no-control strategy but has slightly less effect on the prevalence than the strategy which solely apply full treatment effort with no preventive mechanisms. However, as can be seen in [Fig. 1b](#) the corresponding cost is much higher. If we combine both the preventive controls and the treatment strategies simultaneously, the prevalence comes

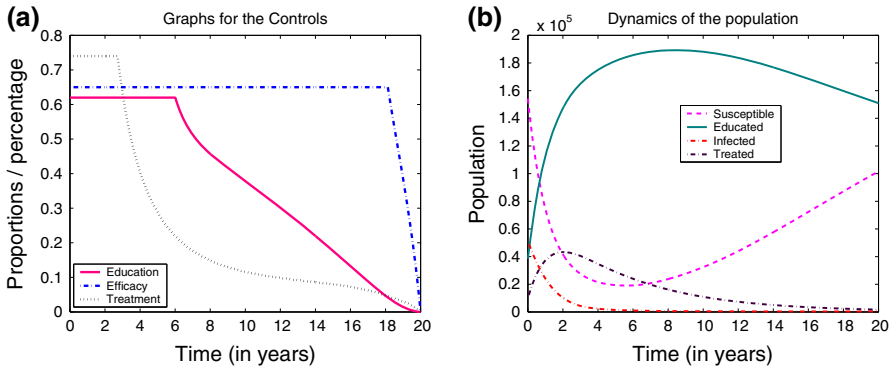


**Fig. 1** a The variation of the prevalence of the disease, when the control parameters vary according to the legend in the graph. b The graph of the marginal cost of the interventions (per day) in various cases of the controls; and c The graph of the incidence of the disease in the population in various cases of the controls; with parameter values are as in Table 1

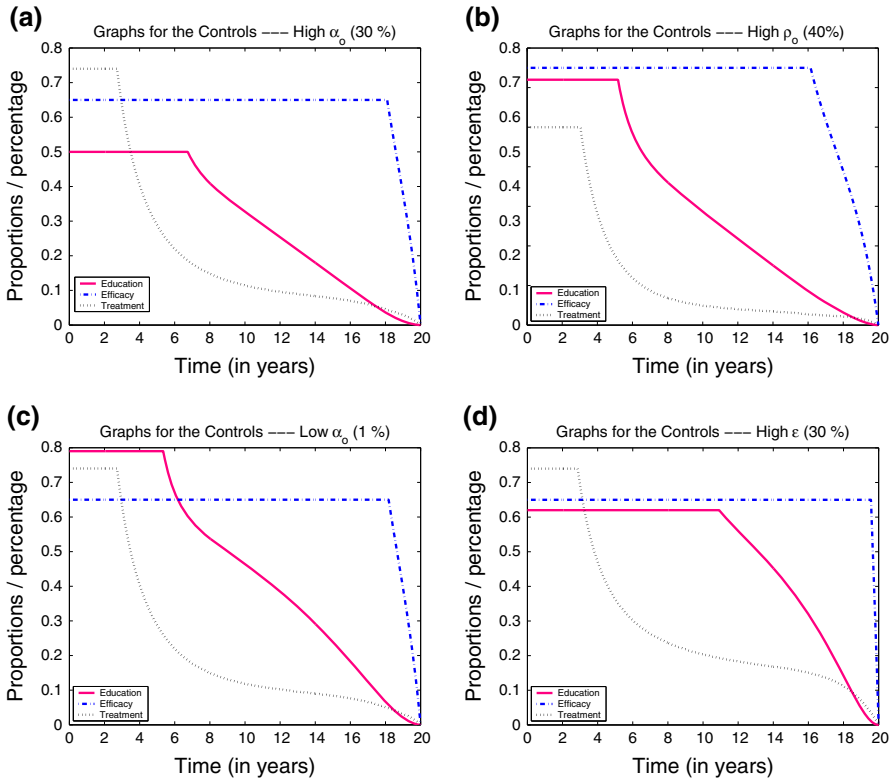
down at a faster rate and the long-term cost is also much better as compared to the other strategies.

In the strategy where all the controls are being used it is optimal to apply all existing resources to each of the control measures at the beginning. But the focus on increasing the rate of recruiting for treatment can be dropped sharply soon-after as compared to the other efforts. On the other hand, educating people and convincing them to participate in self protective schemes must be continued further with slow decrease in the intensity, as can be seen in Fig. 2a. This will help the incidence of the disease not to be high (see Fig. 1c).

The optimality system is not highly sensitive to the choice of the initial value parameters, like  $\alpha_0$  and  $\rho_0$ . However, as can be seen in Fig. 3, the plan for treatment is responsive to changes in the values of  $\alpha_0$  and  $\rho_0$  as well as the change in  $\epsilon$ . The other



**Fig. 2** **a** The graph of optimal control values when all controls are employed simultaneously. **b** The variation of sizes of the four classes of the population when all controls are employed simultaneously



**Fig. 3** The graph of the optimal controls when some of the parameters changed from what is given in Table 1. **a** Shows the graphs of the optimal controls when  $\alpha_o = 0.30$ , **b** shows the graphs of the optimal controls when  $\rho_o = 0.40$ , **c** shows the graphs of the optimal controls when  $\alpha_o = 0.01$ , and **d** shows the graphs of the optimal controls when  $\epsilon = 0.30$

control values show only slight changes in their appearance to adjust the corresponding situations. All the remaining indicators, such as prevalence, incidence, and marginal cost values have shown almost no response to these changes.

In all the cases one does not need to continue on the maximum rate of treatment to obtain the maximum decrease in the prevalence of the disease. After a few years of applying treatment on full possible scale, the recruitment rate for treatment can be reduced to a lower rate and one needs to focus on preventive mechanisms. However, when the value of  $\epsilon$ , the factor of infectiousness of treated individuals, is very high (i.e. when treatment reduces the infectiousness level by only  $(1 - \epsilon)\%$ ) keeping the treatment at higher rate is optimal.

Generally, to get the best result the control on the enhancement of the efficacy level of the protective mechanisms should be given more emphasis with continued education of the population on existing preventive mechanisms coming at the next level in priority. With all the controls employed simultaneously in an optimal way, the prevalence can possibly drop to less than 5 % within 10 years time. The graph for the corresponding dynamics of the disease is given in Fig. 2c.

## 5 Discussion and conclusion

In this paper we derived and analysed a deterministic SIR type model for infectious disease dynamics that takes behaviour modification of the population in to account. We have shown that our mathematical model can portray the way how the population reacts to an increase in prevalence in the course of an outbreak and how one can plan medical treatment to control disease epidemics. In the analysis, it is indicated that behaviour modification by society plays an important role in controlling an epidemic, even when some pharmaceutical treatments are being given to the infected ones. The optimal control theory is used to explain dominance of the behavioural change interventions as compared to treatments. The optimality system also proposes the cost effective way of controlling a disease when behaviour modification and treatment are being implemented on the population level at the same time.

In planning to combat disease epidemics that have no curing medicines, it is important to get the best combination of behaviour change efforts for those who are susceptible and treatments for the infected once. In practice the investment on producing various alternatives of self-protective measures is not that high. However, our simulation shows that it is more effective and cheaper to make more emphasis on such control mechanisms. If the model parameters are estimated well to fit field data the model can predict well the optimal combination of efforts in controlling diseases in human population.

In particular, the dynamics of diseases like HIV can be well presented using this model. Behaviour change efforts, with higher emphasis given to producing alternative self-protective mechanisms, can result in dramatically reducing the prevalence level of the disease. Recently, it has been shown mathematically by [Granich et al. \(2009\)](#) that the prevalence of HIV can be reduced to less than 1 % within 50 years if universal access to ART is implemented. However, our model shows that if it is supplemented by an effective behaviour modification strategy, a 90 % coverage for the first 3 to

5 years and around 50 % coverage for the remaining times of treatment would attain the same result nearly within 18 years.

In using this model for particular disease type in practical applications, one need to take into account the time lag between the determination of the actual value the prevalence of the disease and the reaction of the population. But if the planning time is long enough (which could be counted in years) the prevalence level could be known within a relatively shorter period of time and people could get information about it soon. Hence, the model can be used without additional modification to plan disease control mechanisms.

The model in this paper assumes that people once convinced to change their behaviour, remain in their status forever. But in practice when the disease is endemic for longer time in the population, some individuals may become negligent and go back to practicing risky behaviour. This aspect is not well addressed here. As an extension to their work one can investigate the impact on the dynamics of a disease of such kind of backward flow as well as the effect of the time lag in calculating the prevalence of the disease.

**Acknowledgments** The most part of this work has been done while both authors were working at African Institute for Mathematical Sciences (AIMS), Cape Town. Moreover, Semu Mitiku Kassa acknowledges with thanks the support he received from the International Science Program (ISP) of Sweden, during the preparation of this manuscript.

## References

- Ahituv A, Hotz VJ, Philipson T (1996) The responsiveness of the demand for conditions to the local prevalence of AIDS. *J Hum Resour* 31:869–897
- Behncke H (2000) Optimal control of deterministic epidemics. *Optim Control Appl Methods* 21:269–285
- Bertrand JE (2004) Diffusion of innovations and HIV / AIDS. *J Health Commun* 9:113–121
- Brauer F, van den Driessche P, Wu J (eds) *Mathematical epidemiology*. Springer, Berlin (2008)
- Campbell GL, Hughes JM (1995) Plague in India: a new warning from an old nemesis. *Ann Intern Med* 122:151–153
- Castillo-Chavez C, Song B (2004) Dynamical models of tuberculosis and their applications. *Math Biosci Eng* 1:361–404
- Chen FH (2004) Rational behavioural response and the transmission of STDs. *Theor Popul Biol* 66:307–316
- Chen FH (2006) On the transmission of HIV with self-protective behaviour and preferred mixing. *Math Biosci* 199:141–159
- Chen FH (2009) Modelling the effect of information quality on risk behaviour change and the transmission of infectious diseases. *Math Biosci* 217:125–133
- Coddington EA, Levinson N (1972) *Theory of ordinary differential equations*, TMH edition. McGraw-Hill, NY
- Donnell D, Baeten JM, Kiarie J, Thomas KK, Stevens W, Cohen CR, McIntyre J, Lingappa JR, Celum C (2010) Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. *Lancet* 375:2092–2098
- Fenichel EP, Castillo-Chavez C, Ceddia MG, Chowell G, Gonzalez Parra PA, Hickling J, Holloway G, Horan R, Morin B, Perrings C, Springborn M, Velazquez L, Villalobos C (2011) Adaptive human behavior in epidemiological models. *PNAS* 108(15):6306–6311
- Ferguson N (2007) Capturing human behaviour. *Nature* 446:733–733. doi:10.1038/446733a
- Fister KR, Lenhart S, McNally JS (1998) Optimizing chemotherapy in an HIV model. *Electron J Differ Equ* 32:1–12
- Funk S, Gilad E, Watkinsen C, Jansen VAA (2009) The spread of awareness and its impact on epidemic outbreaks. *Proc Natl Acad Sci USA* 106:6872–6877

- Funk S, Salathé M, Jansen VAA (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J R Soc*. doi:[10.1098/rsf.2010.0142](https://doi.org/10.1098/rsf.2010.0142)
- Gaff H, Schaefer E (2009) Optimal control applied to vaccination and treatment strategies for various epidemiological models. *Math Biosci Eng* 6:469–492
- Granich RM, Gilks CF, Dye C, De Cock KM (2009) Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. *Lancet* 375:48–57
- Green LW, McAlister AL (1984) Macro-intervention to support health behaviour: some theoretical perspectives and practical reflections. *Health Educ Q* 11:322–339
- Green LW, Ottoson JM, García C, Hiatt RA (2009) Diffusion theory and knowledge dissemination, utilization, and integration in public health. *Annu Rev Pub Health* 30:151–174
- Hackbush WK (1978) A numerical method for solving parabolic equations with opposite orientations. *Computing* 20:229–240
- Hartl RF, Sethi SP, Vickson RG (1995) A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev* 37:181–218
- Hethcote HW (2000) The mathematics of infectious diseases. *SIAM Rev* 42:599–653
- Joshi H, Lenhart S, Albright K, Gibson K (2008) Modelling the effect of information campaigns on the HIV epidemic in Uganda. *Math Biosci Eng* 5(4):757–770
- Kassa SM, Ouhinou A (2011) Epidemiological models with prevalence dependent endogenous self-protection measure. *Math Biosci* 229(1):41–49
- Lau JT, Yang X, Pang E, Tsui HY, Wong E, Wing YK (2005) SARS-related perceptions in Hong Kong. *Emerg Infect Dis* 11:417–424
- Lukes DL (1982) *Differential equations: classical to control*. Academic press, Edinburgh
- Oldenburg B, Glanz K (2008) Diffusion of Innovations. In: Glanz K, Rimer BK, Viswanath K (eds) *Health behaviour and health education*, 4th edn. Wiley, London
- Philipson T (1996) Private vaccination and public health: an empirical examination for US measles. *J Hum Resour* 31:611–630
- Pontryagin LS, Boltyanskii VT, Gamkrelidze RV, Mishchenko EF (1962) *The mathematical theory of optimal processes*. Wiley, London
- Rogers EM (1983) *Diffusion of innovations*, 3rd edn. The Free press, New York
- Turner CF, Miller HG, Moses LE (eds) (1989) *AIDS, sexual behaviour, and intravenous drug use*. National Academy of Sciences, Washington, DC
- Vardavas R, Blower S (2007) The emergence of HIV transmitted resistance in Botswana: when will the WHO detection threshold be exceeded? *PLoS ONE* 2(1):e152. doi:[10.1371/journal.pone.0000152](https://doi.org/10.1371/journal.pone.0000152)
- van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibrium for compartmental models of disease transmission. *Math Biosci* 180:29–48
- World Health Organization (2005) Summary country profile for HIV/AIDS treatment scale-up. Botswana. [http://www.who.int/hiv/HIVCP\\_BWA.pdf](http://www.who.int/hiv/HIVCP_BWA.pdf). Retried 19 April 2012